NO. 81644-1

SUPREME COURT OF THE STATE OF WASHINGTON

In Re the Detention of:	STATEMENT OF
	ADDITIONAL AUTHORITY
DAVID MCCUISTION,	
Petitioner-Appellant.	

As permitted by RAP 10.8, the State of Washington, respondent, submits the following additional authority:

Grant Duwe et al., *The Impact of Prison-Based Treatment on Sex Offender Recidivism*, Sex Abuse OnlineFirst, June 16, 2009, at 1, 12, 18:

"[P]articipating in treatment significantly reduced the hazard ratio for rearrest by 27% for sexual recidivism...These findings are consistent with the growing body of research supporting the effectiveness of cognitive-behavioral treatment for sex offenders."

"Since the mid-1990s, however, meta-analyses of the treatment literature have, with a few notable exceptions, found lower sexual recidivism rates for treated sex offenders in comparison with untreated offenders. Among the meta-analyses that have found a treatment effect, the rate of sexual reoffense has been between 5 and 10 percentage points less for those who participated in treatment..."

"[T]reated offenders had lower reoffense rates than untreated offenders for each of the three types of recidivism − sexual, violent, and general. Not surprisingly, the best recidivism outcomes were found for offenders who completed treatment or successfully participated until their release."

"Sex offenders who participated in treatment recidivated less often and more slowly than untreated offenders."

"The average sexual recidivism rate was 27% lower for treated offenders than for untreated offenders..."

R. Karl Hanson et al., *First Report of the Collaborative Outcome Data Project on the Effectiveness of Psychological Treatment for Sex Offenders*, Sexual Abuse: A Journal of Research and Treatment, Vol. 14, No. 2, April 2002, at 169:

> "This meta-analytic review examined the effectiveness of psychological treatment for sex offenders by summarizing data from 43 studies (combined n=9,454). Averaged across all studies, the sexual offence recidivism rate was lower for the treatment groups (12.3%) than the comparison groups (16.8%, 38 studies, unweighted average)."

Friedrich Losel et al., *The Effectiveness of Treatment for Sexual Offenders: A Comprehensive Meta-Analysis*, Journal of Experimental Criminology, 2005, at 135, 138:

> "The mean rate of sexual recidivism is 11.1% in TG's (treatment group) and 17.5% in CGs (comparison group). At first glance, this absolute difference of a little more than 6 percentage points may seem small. However, when the low base rate of sexual recidivism is taken into account, this is equivalent to a reduction of nearly 37%."

> "Overall, there is evidence for a positive effect of sexual offender treatment. Cognitive-behavioral and hormonal treatment are most promising."

Kathy Gookin, *Comparison of State Laws Authorizing Involuntary Commitment of Sexually Violent Predators: 2006 Update, Revised*, Washington State Institute for Public Policy, August 2007, at 4:

> Exhibit 2, 2006 Program Overview, Reasons for Discharges and Releases of People Committed... Washington . . . Program Staff Recommendation... 12. [Twelve individuals released from Special Commitment Center with recommendation of treatment program staff].

Diana Hefley, *Sex Offender Released After Treatment*, The Everett Daily Herald, July 19, 2007, at 1:

> "While locked up Paschke underwent sex offender treatment. He is the first sexually violent predator who completed the

entire program . . . A Snohomish County judge signed an order for his release Wednesday."

Keith Eldridge, *One of State's Most Notorious Sex Offenders Set Free,* KOMO News, August 15, 2008, at 1-2, www.komonews.com/news/local /27037919.html:

> "One of the state's most notorious sex offenders has just been set free. John Mathers won his freedom after spending 12 years at the state's Special Commitment Center on McNeil Island. On Friday Mathers stood before a judge one last time just minutes before regaining his freedom. Psychiatrists concluded the 56 year old has completed his sex offender treatment and is ready for release back into the community."

Court Frees McNeil Island Sex Felon, The Seattle Post-Intelligencer, August 21, 2008, at 1:

> "A Pierce County Superior court judge ordered 56-year-old John Henry Mathers released on Friday because he has completed a treatment program."

Richard Thompson, *Man Who Admitted Raping 22 Women To Be Released From Program,* KIRO 7 Eyewitness News, Aug. 6, 2009, http://www.kirotv.com/crime/20310854 /detail.html.

> "DSHS said Cherry has completed all of his treatment and evaluations show he can no longer be labeled a "sexually violent predator" and must be released without conditions."

RESPECTFULLY SUBMITTED this 13th day of October, 2009,

ROBERT M. MCKENNA
Attorney General

TODD BOWERS, WSBA #25274
Senior Counsel
Attorneys for Respondent
(206) 389-2028

3

The Impact of Prison-Based Treatment on Sex Offender Recidivism

Evidence From Minnesota

Grant Duwe
Minnesota Department of Corrections, St. Paul
Robin A. Goldman
Minnesota Department of Corrections, Lino Lakes

Using a retrospective quasi-experimental design, this study evaluates the effectiveness of prison-based treatment by examining recidivism outcomes among 2,040 sex offenders released from Minnesota prisons between 1990 and 2003 (average follow-up period of 9.3 years). To reduce observed selection bias, the authors used propensity score matching to create a comparison group of 1,020 untreated sex offenders who were not significantly different from the 1,020 treated offenders. In addition, intent-to-treat analyses and the Rosenbaum bounds method were used to test the sensitivity of the findings to treatment refuser and unobserved selection bias. Results from the Cox regression analyses revealed that participating in treatment significantly reduced the hazard ratio for rearrest by 27% for sexual recidivism, 18% for violent recidivism, and 12% for general recidivism. These findings are consistent with the growing body of research supporting the effectiveness of cognitive–behavioral treatment for sex offenders.

Keywords: sex offender; treatment; prison; recidivism; propensity score matching

Over the past 50 years, dozens of studies from a number of countries have examined whether sex offender treatment reduces recidivism. Reviews of the earliest studies drew pessimistic conclusions about the effectiveness of treatment. For example, in their 1989 review of the treatment literature, Furby, Weinrott, and Blackshaw argued that, due to methodological shortcomings, there was insufficient evidence to support the notion that treatment decreases sex offender recidivism. Several years later, Quinsey, Harris, Rice, and Lalumiere (1993) reached a similar conclusion in their review of existing treatment studies.

Since the mid-1990s, however, meta-analyses of the treatment literature have, with a few notable exceptions (Kenworthy, Adams, Brooks-Gordon, & Fenton, 2004; Rice & Harris, 2003), found lower sexual recidivism rates for treated sex offenders in comparison with untreated offenders (Alexander, 1999; Gallagher, Wilson, Hirschfield, Coggeshall, & MacKenzie, 1999; Hall, 1995; Hanson et al., 2002; Lösel & Schmucker, 2005). Among the meta-analyses that have found a treatment effect, the rate of sexual reoffense has been between 5 and 10 percentage points less for those who participated in treatment, resulting in a mean effect size (Cohen's d) ranging from .12 to .47. The evidence from these studies further indicates that cognitive–behavioral techniques with relapse prevention components have, by and large, been found to be the most effective in reducing recidivism.

Despite the generally positive findings from the meta-analytical reviews, it is nevertheless true that most of the existing treatment studies have lacked methodological rigor. In the Lösel and Schmucker (2005) study, which is the most comprehensive meta-analysis to date, 80 comparisons (69 studies) between treated and untreated sex offenders were examined. Of these comparisons, only 6 (7%) used a randomized experimental design—most notably, the research by Marques and colleagues (Marques, 1999; Marques, Day, Nelson, & West, 1994; Marques, Wiederanders, Day, Nelson & von Ommeren, 2005)—whereas 7 (9%) used individual matching or statistical control in an effort to achieve equivalence between the treatment and comparison groups. Instead, most treatment studies have used either nonequivalent comparison groups (60%) or research designs in which equivalence was assumed between the treated and untreated groups (24%).

Given the relatively large percentage (84%) of studies that have not used random assignment or matching techniques, selection bias is, as some have pointed out (Harkins & Beech, 2007; Jones, Pelissier, & Klein-Saffran, 2006; Rice & Harris, 2003), a problem that has plagued the sex offender treatment literature. In evaluations of treatment effectiveness, selection bias refers to differences—both observable and unobservable—between the treated and untreated groups that make it difficult to determine whether the observed effects are due to the treatment itself or to the different group compositions. Therefore, although previous evaluations have found that recidivism rates are generally reduced for sex offenders who participate in treatment, this effect may not necessarily be due to the treatment itself but rather to other differences between treated and untreated offenders.

In addition to selection bias, the vast majority of existing studies share a number of limitations. For example, of the studies reviewed by Lösel and Schmucker (2005), only 11 (13%) had a sample size in excess of 500. Moreover, the follow-up periods for many studies have been relatively short, as only one fourth of those examined by Lösel and Schmucker tracked offenders for more than 7 years.

Present Study

In evaluating the effectiveness of sex offender treatment in Minnesota prisons, this study does not use a randomized experimental design. Furthermore, due to a

lack of available data, it does not control for the possible impact that postrelease participation in community-based treatment may have on reoffending. Despite these limitations, however, the present study contains a number of strengths that have been lacking from most prior treatment studies. First, as discussed later in more detail, we used propensity score matching (PSM) to individually match treated and untreated sex offenders. In doing so, this study minimizes the threat of selection bias by creating a comparison group whose probability of entering treatment was similar to that of the treatment group. Second, in addition to being one of the first studies in the sex offender treatment literature to use PSM (Caldwell, Skeem, Salekin, & Van Rybroek, 2006; Skeem, Monahan, & Mulvey, 2002), this study further controls for rival causal factors by analyzing the data with Cox regression, which is widely regarded as the most appropriate multivariate statistical technique for recidivism analyses. Third, by comparing 1,020 treated sex offenders with a matched group of 1,020 untreated sex offenders, the sample size used for this study ($N = 2,040$) is one of the larger sex offender treatment studies to date. Fourth, to gain a more precise assessment of the effectiveness of treatment, we used multiple measures of treatment participation and criminal recidivism. Finally, because recidivism data were collected on the 2,040 sex offenders through the end of 2006, the average follow-up period for these offenders was 9.3 years. This study thus provides a robust assessment of treatment effectiveness by tracking offenders over a relatively lengthy period of time.

In examining prison-based treatment in Minnesota, we attempt to address several questions central to the sex offender treatment literature. First, does treatment participation reduce offender recidivism? Second, what effect does treatment outcome (i.e., dropout, complete, successfully participate until release, etc.) have on reoffending? Finally, are there certain types of sex offenders for whom treatment is more effective?

In the following section, we describe the provision of sex offender treatment within the Minnesota Department of Corrections (MNDOC). The data and methods used in this study are then discussed, followed by a presentation of the results. We conclude by discussing the implications of the findings for the sex offender treatment literature.

Program Description: Sex Offender Treatment in the MNDOC

In 1978, the MNDOC began providing sex offender treatment to incarcerated adult men when it opened the Transitional Sex Offender Treatment Program (SOTP)—a 30-bed program for offenders preparing to return to the community—at the Minnesota Correctional Facility (MCF) at Lino Lakes. Consistent with the name of the program, offenders were treated in the last year of their incarceration and the emphasis was on preparation for release. As a component of the program design,

services were continued for men postrelease in a half-way house setting in conjunction with the staff of the half-way house.

In 1983, a second prison-based sex offender treatment program was established at MCF–Oak Park Heights, Minnesota's lone maximum custody facility. In addition to providing sex offender treatment, this 52-bed program treated chemically dependent inmates as well as those with dual diagnoses (i.e., needing treatment for both chemical dependency and sexual offending). In 1994, this program relocated to the medium custody facility, MCF–Lino Lakes, and eventually integrated with the original SOTP, which had grown and evolved since 1978.

In 1991, the Sexual Education and Evaluation Center, a small (36-bed) sex offender treatment program opened at MCF–Stillwater. Psychoeducation and therapy groups were the primary services offered by this program, which was designed to provide a short-term but intensive treatment experience. Three years later, a 50-bed sex offender treatment program opened at MCF–Moose Lake. Designed to provide long-term intensive sex offender treatment, this program also offered some specialized groups for offenders with intellectual or cognitive deficits.

Due to prison population growth and increased sentence lengths for sex offenders, the size of the program at the MCF–Lino Lakes increased to 110 beds by 1997 and then further expanded to 150, when the Sexual Education and Evaluation Center program was transferred to MCF–Lino Lakes and integrated with the program at that site. In 2000, it expanded once again with the transfer and integration of the MCF–Moose Lake program to MCF–Lino Lakes.

Over the past three decades, sex offender treatment programming at MCF–Lino Lakes evolved to keep pace with changing practices in the field, while attempting to maintain the most unique and seemingly valuable components of each of the programs that were melded into the program that exists today. For example, the chemical dependency (CD) treatment component was maintained. In doing so, the SOTP addressed the chemical abuse issues, and their relationship to sexual offending, in a more integrated manner than would likely occur in a standalone CD treatment program. In addition, psychoeducation classes and therapy designed to accommodate the needs of the offender with intellectual or cognitive functioning limitations were maintained. Moreover, the number and variety of psychoeducation classes designed to impart information and enhance skill building were expanded at the SOTP. Finally, an emphasis on preparation for release was also maintained.

Using a cognitive–behavioral framework, the SOTP attempted to provide long-term intensive sex offender and CD treatment consistent with a risk-needs-responsivity model. To be eligible for treatment, offenders had to have at least 9 months to serve in prison. Moreover, offenders who minimized their offenses (as described in official documents) were eligible to enter treatment in the MNDOC, whereas those who completely denied committing a sexual offense were not eligible. Given the fact that treatment capacity did not keep pace with the overall growth in inmate population, the SOTP attempted to target moderate- to high-risk sex offenders for treatment.

Offenders considered to be lower risk were less likely to be admitted to sex offender treatment programming while incarcerated. However, offenders required to enter treatment but who were unable to do so while incarcerated were recommended to participate in community-based treatment at the time of release.

Under the current process, offenders are prioritized for treatment primarily on the basis of their scores from the following actuarial instruments: Static-99, Rapid Risk Assessment for Sex Offense Recidivism, and Minnesota Sex Offender Screening Tool–Revised. Earlier attempts, however, to identify and prioritize cases for treatment were based on more primitive tools such as the Public Risk Monitoring (PRM) criteria, which were developed by MNDOC staff. Offenders who met the PRM criteria, which were never formally validated on the sex offender population, were directed to participate in treatment programming. As shown later, the PRM criteria did not appear to be very effective in distinguishing offenders on the basis of recidivism risk. Indeed, the recidivism risk (as reflected by the risk score measure developed for this study) was not significantly different between offenders who were offered treatment and those who were not.

After receiving a treatment directive, offenders had the right to refuse treatment. There were consequences, however, for those who exercised this right. In particular, offenders who failed to comply with a treatment directive had their wages frozen and were subject to extended incarceration disciplinary time that lengthened their stay in prison. In addition, treatment participation and outcome is an item on the Minnesota Sex Offender Screening Tool–Revised, which has been used to guide decisions regarding community notification levels and civil commitment referral. Therefore, even though offenders can refuse the directive to enter treatment, the "carrot-and-stick" approach used by the MNDOC likely motivated many offenders to enter treatment programming who might have otherwise opted not to do so if the choice were entirely voluntary.

SOTP participants were housed in two adjacent wings of a larger living unit. This arrangement allowed for some movement across the wings in the living unit but no movement either to, or from, the other wings where the general population inmates were housed. The living units operated within a modified therapeutic milieu with clear living unit/program structure and rules, and there was an expectation that inmates will support and hold one another accountable throughout their day. The SOTP was not an entirely closed living unit, as there was some interaction with the general inmate population during movement and activities such as dining, religious services, educational programming, and so on. A 30-day assessment and orientation phase, individual and group therapy, and psychoeducational programming were held next to the living unit in two buildings, where the offices of clinical staff were also located. Although individualized treatment plans varied widely among treatment participants, the average dosage consisted of approximately 10 to 15 hours of direct staff facilitated services per week for a duration that often ranged from 1 to 3 years.

In 1999, the MNDOC implemented legislatively mandated rules for residential sex offender treatment that require programs to meet specified staff training and ratios, group and class size limitations, have a written theory-based treatment protocol, and demonstrate that that they apply this to their assessment, treatment, and therapeutic milieu. The program was audited on a biannual basis. Given the state of research and practice in the field of sex offender treatment, the SOTP continues to be a work in progress. The following section describes the main components of the program.

SOTP Components

Assessment. This 30-day phase of the program includes psychological testing, completion of assignments to facilitate the assessment of treatment needs, a review of offending history and offense dynamics, and use of lecture, discussion, and video-tapes to provide information on treatment participation and expectations, defenses and denial, sexual assault dynamics, victim impact, CD, and so on. A clinical interview, collateral information, client observations, and test results were used to develop a written psychosexual assessment report and an individualized treatment plan.

Therapy. Following the assessment phase of the program, inmates participated in an average of 6 hours per week of staff-facilitated group therapy sessions. Therapy groups specific to the needs of the inmate with cognitive/intellectual limitations were provided. Additional individual therapy was offered based on the needs of the inmate and the availability of staff. Therapy was provided in progressive phases and included transitional programming and aftercare. Ongoing therapy and postrelease programming in the community was provided under contract with, or through, grants from the MNDOC to private agencies.

CD treatment. Offenders entering the MNDOC were formally screened, assessed, and diagnosed for chemical abuse or CD. Treatment directives were provided contingent on the outcome of these assessments. For those sex offenders needing treatment for alcohol or drug dependency, CD treatment was typically provided following completion of the Assessment Phase in the SOTP.

Family/support person education. To prepare offenders for their return to the community, while also helping them reach specific treatment goals, education sessions were facilitated between program participants and members of their family and/or support system. These sessions were used to provide clarification about the nature and impact of their offending, to inform support persons about the risk for reoffense, and to identify response strategies for the offender and the support person.

Psychoeducational programming. Program members participated in psychoeducational programming, which varied according to the offender's individualized

treatment plan. Psychoeducational classes were typically provided for 1.5-hour sessions 3 to 4 times per week in 12-week (quarter) sessions. Each quarter, program participants were each enrolled in one or two classes, which included Emotions Management, Alcohol and Drug Education, Cognitive Restructuring and Criminal Thinking, Sexuality Education, Sexual Assault Dynamics, Reoffense Prevention, Victim Empathy, Personal Victimization, Grief and Loss, Morals and Values, Sexual Behaviors, and Transitional Curriculum. Classes specific to the needs of the inmate with cognitive limitations included modifications of some of the classes above.

In addition to these psychoeducational classes, offenders were assigned, per their individual treatment plan, to participate in a parenting class provided in the institution under a contract with a nonprofit agency. Offenders in the program also participated in additional educational, prerelease, and transitional planning (housing, employment, transportation, etc.), which drew heavily on resources from the community.

Support groups. On a weekly basis, offenders attended additional support groups such as Alcoholics Anonymous, Narcotics Anonymous, and Sex Abusers Anonymous meetings held in the institution. Each of these groups met for 1.5 hours per week and were monitored, but not facilitated, by program staff.

Community meetings. Inmates met weekly in a large group with other members of their living unit to address general housekeeping issues, community milieu, and to provide support to one another as a community.

Data and Methodology

To determine whether treatment provided within the MNDOC has had an impact on sex offender recidivism, we used a retrospective quasi-experimental design. That is, we evaluated the effectiveness of sex offender treatment by comparing recidivism outcomes between treated offenders and a matched comparison group of untreated offenders who were released between 1990 and 2003. During this 14-year period, there were 3,440 sex offenders who were released from Minnesota prisons.

Of these offenders, 1,493 (43%) participated in prison-based treatment prior to their release from prison. Of the remaining 1,947 offenders, 105 refused to enter treatment whereas the other 1,842 offenders were not given the opportunity to participate. Because the 105 treatment refusers did not participate in treatment, we removed these offenders from the study so as not to bias the results from the statistical analyses. Before doing so, however, we attempted to remove an additional source of bias by using PSM to identify a comparison group of 105 offenders from the pool of untreated offenders ($N = 1,842$) who were not offered treatment.

Propensity Score Matching

PSM is a method that estimates the conditional probability of selection to a particular treatment or group given a vector of observed covariates (Rosenbaum & Rubin, 1984). The predicted probability of selection, or propensity score, is typically generated by estimating a logistic regression model in which assignment ($0 = no$ $assignment$; $1 = assignment$) is the dependent variable whereas the predictor variables consist of those that theoretically have an impact on the selection process. Shadish, Cook, and Campbell (2002) note that unless the predictors are unrelated to the outcome variable or are not proper covariates, they should be included in the propensity score model even if they are only weakly associated with the outcome (recidivism for this study). Once estimated, the propensity scores are then used to match individuals who entered treatment (or refused to enter treatment) with those who did not. Thus, one of the main advantages with using PSM is that it can simultaneously "balance" multiple covariates on the basis of a single composite score. Although there are a number of different matching methods available, we used a "greedy" matching procedure that used a without-replacement method in which treated offenders were matched to untreated offenders who had the closest propensity score (i.e., "nearest neighbor") within a caliper (i.e., range of propensity scores) of .10 (defined in terms of probabilities).

In matching untreated offenders with treated offenders on the conditional probability of entering treatment, PSM reduces selection bias by creating a counterfactual estimate of what would have happened to the treated offenders had they not participated in treatment. PSM is not without its limitations, however. First, and foremost, because propensity scores are based on observed covariates, PSM is not robust against "hidden bias" from unmeasured variables that are associated with both the assignment to treatment and the outcome variable. Second, there must be substantial overlap among propensity scores between the two groups for PSM to be effective (Shadish et al., 2002); otherwise, the matching process will yield incomplete or inexact matches. Finally, as Rubin (1997) points out, PSM tends to work best with large samples.

Although somewhat limited by the data available to us, we attempted to address potential concerns over unobserved bias by including as many theoretically relevant covariates (17) as possible in our propensity score models. More important, however, we conducted Rosenbaum bounds sensitivity analyses to evaluate the extent to which the treatment effects obtained are robust to the possibility of hidden bias. In addition, we later demonstrate that there was substantial overlap in propensity scores between the treated and untreated offenders. Furthermore, we addressed the sample size limitation by assembling a relatively large number of cases ($N = 3,440$) on which to conduct the propensity score analyses.

Matching treatment refusers and nonrefusers. In an effort to minimize the bias resulting from treatment refusers, we attempted to identify a comparison group of

untreated offenders who were not offered treatment in order to remove these offenders from the comparison group pool. We computed propensity scores for the 105 treatment refusers and the 1,842 untreated offenders by estimating a logistic regression model in which the dependent variable was refusal of treatment (i.e., the 105 treatment refusers were assigned a value of "1," whereas the 1,842 untreated offenders in the comparison group pool received a value of "0"). The predictors were the 17 control variables, which are described later, that were used in the statistical analyses. After obtaining propensity scores on the 1,947 offenders, we used the greedy matching procedure to match 105 untreated offenders not offered treatment with the 105 treatment refusers.

In Table 1, we present the covariate, propensity score, and recidivism outcome means for both groups prior to matching ("total") and after matching ("matched"). In addition to tests of statistical significance ("t test p value"), we provide a measure ("Bias") developed by Rosenbaum and Rubin (1985) that quantifies the amount of bias between the treatment and control samples (i.e., standardized mean difference between samples),

$$\text{Bias} = \frac{100(\overline{X}_t - \overline{X}_c)}{\sqrt{(S_t^2 + S_c^2)/2}}$$

where $\overline{X}_t$ and S_T^2 represent the sample mean and variance for the treated offenders and $\overline{X}_c$ and S_c^2 represent the sample mean and variance for the untreated offenders. If the value of this statistic exceeds 20, the covariate is considered to be unbalanced (Rosenbaum & Rubin, 1985).

As shown in Table 1, the matching procedure reduced the bias in propensity scores between treatment refusers and those not offered treatment by 99%. Whereas the p value was .00 in the unmatched sample, it was .95 in the matched sample. Although risk score was not used as a predictor in the logistic regression analysis, we also present the means for this variable to illustrate the differences between the two groups before and after matching. In the unmatched sample, there were nine covariates that were significantly imbalanced (i.e., the difference between the treatment refusers and those not offered treatment was significant at the .05 level and the bias values exceeded 20). But in the matched sample, covariate balance was achieved insofar as there were no covariates with bias values greater than 20 or with significant differences between the treatment refusers and those not given a treatment opportunity. Just as important, when examining the outcome data for these two groups of offenders within the unmatched sample, we see that treatment refusers had significantly higher rates of sexual and violent recidivism. In the matched sample, however, recidivism outcomes were not significantly different between the two groups. Along with the 105 treatment refusers, we removed the 105 matched offenders not offered treatment from the remaining analyses. In doing so, we reduced the number of untreated offenders in the comparison group pool by 210 from 1,947 to 1,737.

Table 1
Propensity Score Matching and Covariate Balance for Refusers

Variable	Sample	Refusers Mean	Nonrefusers Mean	Bias (%)	Bias Reduction	t Test p Value
Propensity score	Total	0.11	0.05	53.25		.00
	Matched	0.11	0.11	0.76	−98.58%	.95
Minority	Total	38.10%	37.79%	0.52		.95
	Matched	38.10%	39.05%	1.59	205.38%	.89
Age at release (years)	Total	32.89	33.04	1.08		.89
	Matched	32.89	32.84	0.37	−65.72%	.98
Metro	Total	49.52%	45.44%	6.66		.41
	Matched	49.52%	59.05%	15.58	134.05%	.17
Prior sex crimes	Total	41.90%	20.20%	37.98		.00
	Matched	41.90%	44.76%	4.70	−87.64%	.68
Prior felony	Total	68.57%	63.79%	8.28		.32
	Matched	68.57%	66.67%	3.31	−60.04%	.77
Stranger	Total	10.48%	9.50%	2.64		.74
	Matched	10.48%	12.38%	4.91	85.96%	.67
Acquaintance	Total	66.67%	60.42%	10.66		.20
	Matched	66.67%	64.76%	3.28	−69.24%	.77
Adult female	Total	16.19%	18.78%	5.61		.51
	Matched	16.19%	14.29%	4.26	−24.02%	.70
Male child	Total	3.81%	5.54%	6.89		.45
	Matched	3.81%	1.90%	8.86	28.51%	.41
Length of stay (months)	Total	29.13	23.65	15.43		.01
	Matched	29.13	29.24	0.28	−98.18%	.98
Discipline	Total	1.38	1.45	2.30		.82
	Matched	1.38	1.13	10.90	374.76%	.32
Supervision (months)	Total	50.58	34.17	42.00		.00
	Matched	50.58	43.33	18.00	−57.13%	.11
Intensive supervised release	Total	32.38%	18.19%	26.10		.00
	Matched	32.38%	24.76%	13.58	−47.99%	.22
Supervised release	Total	61.90%	79.59%	31.30		.00
	Matched	61.90%	70.48%	14.65	−53.19%	.19
Supervised release revocations	Total	1.34	0.76	34.69		.00
	Matched	1.34	1.36	0.89	−97.42%	.93
Community notification	Total	14.30%	2.40%	32.33		.00
	Matched	14.30%	14.30%	0.00	−100.00%	1.00
Release year	Total	1998.61	1996.93	32.13		.00
	Matched	1998.61	1997.76	16.02	−50.14%	.16
Risk score	Total	4.39	3.95	19.80		.02
	Matched	4.39	4.32	3.16	−84.05%	.78
	Total					
Total recidivism	Matched					
Sex rearrest	Total	41.0%	16.1%			.00
	Matched	41.0%	36.2%			.48
Violent rearrest	Total	55.2%	33.6%			.00
	Matched	55.2%	47.6%			.27
Any rearrest	Total	66.7%	59.8%			.16
	Matched	66.7%	69.5%			.66

Note: Total refusers, $N = 105$; total nonrefusers, $N = 1,842$; matched refusers, $N = 105$; matched nonrefusers, $N = 105$.

Table 2
Logistic Regression Model for Assignment to Treatment

Predictors	Coefficient	Standard Error
Minority	−0.292**	0.088
Age at release (years)	9.0E-4*	3.9E-4
Metro	0.290**	0.081
Prior sex crime convictions	0.789**	0.095
Prior felony convictions	−0.348**	0.082
Stranger victims	−0.048	0.153
Acquaintance victims	−0.039	0.089
Adult Female victims	−0.093	0.106
Male child victims	0.145	0.160
Length of stay (months)	0.021**	2.0E-4
Discipline	−0.076**	0.019
Supervision (months)	6.5E-4**	1.6E-4
Intensive supervised release	0.980**	0.368
Supervised release	0.902*	0.358
Supervised release revocations	0.160**	0.033
Community notification	0.703**	0.238
Release year	−0.017	0.014
Constant	31.997	27.806
N	3,230	
Log-likelihood	3929.875	
Nagelkerke R^2	.202	

*$p < .05$. **$p < .01$.

Matching treated and untreated sex offenders. Similar to the approach described above with treatment refusers, we calculated propensity scores for the 1,493 treated offenders and the 1,737 untreated offenders by estimating a logistic regression model in which the dependent variable was participation in prison-based treatment (i.e., the 1,493 group offenders were assigned a value of "1," whereas the 1,737 offenders in the comparison group pool received a value of "0"). The predictors were the 17 control variables (excluding risk score) used in the statistical analyses (see Table 2). As shown in Figure 1, there was substantial overlap in propensity scores between the treated and untreated offenders, even though the difference in mean propensity score was statistically significant at the .01 level (see Table 3).

After obtaining propensity scores for the 3,230 offenders, we used the greedy matching procedure to match the untreated offenders with the treated offenders. Because the matching process is often a trade-off between the size of the bias reduction and the proportion of cases that can be matched (DiPrete & Gangl, 2004), especially in situations like this one where the treated offenders ($N = 1,493$) accounted for nearly half (46%) of the offenders ($N = 3,230$), we were unable to
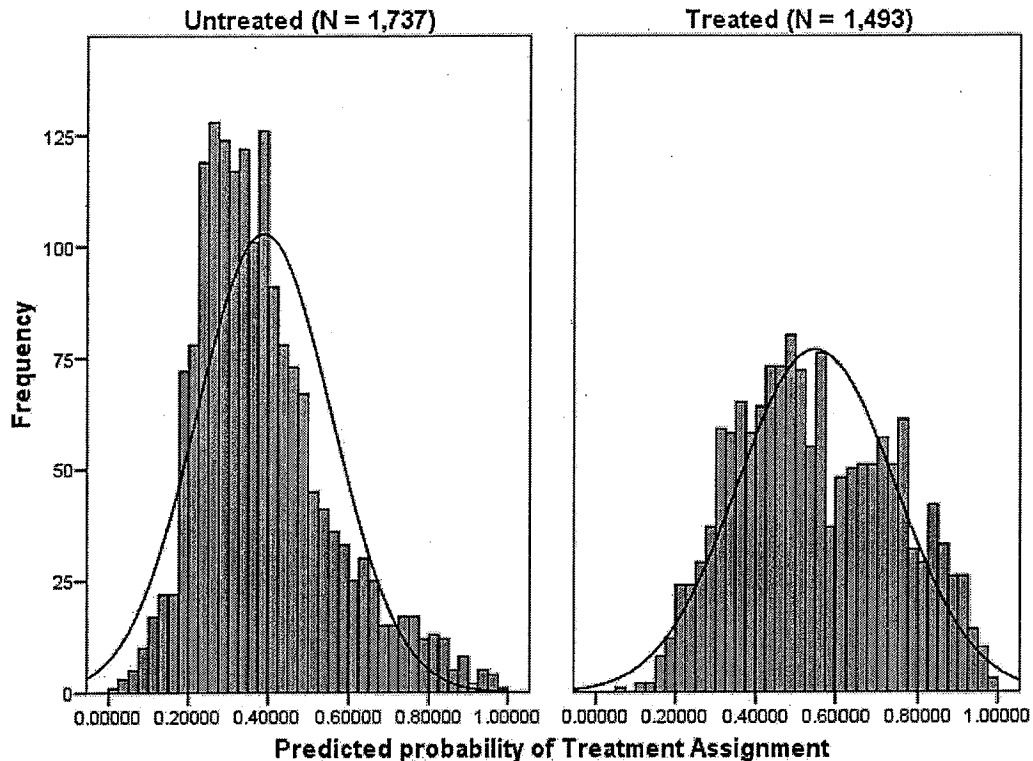
Table 3
Propensity Score Matching and Covariate Balance for Treatment

Variable	Sample	Treated Mean	SD	Untreated Mean	SD	Bias (%)	Bias Reduction	t Test p Value
Propensity score	Total	0.55	0.19	0.39	0.17	70.31		.00
	Matched	0.47	0.17	0.47	0.17	1.19	−98.30%	.74
Minority	Total	33.09%	0.47	37.71%	0.48	7.94		.01
	Matched	35.20%	0.48	35.10%	0.48	0.15	−98.06%	.96
Age at release	Total	36.10	10.48	33.05	11.06	23.33		.00
(years)	Matched	34.88	10.23	34.94	11.80	0.52	−97.79%	.90
Metro	Total	51.44%	0.50	44.62%	0.50	11.16		.00
	Matched	47.94%	0.50	48.53%	0.50	0.84	−92.48%	.79
Prior sex crimes	Total	39.38%	0.49	18.71%	0.39	36.84		.00
	Matched	27.55%	0.45	28.14%	0.45	1.01	−97.27%	.77
Prior felony	Total	54.39%	0.50	63.62%	0.48	15.29		.00
	Matched	58.33%	0.49	56.96%	0.50	1.88	−87.72%	.53
Stranger	Total	11.72%	0.32	9.33%	0.29	6.27		.03
	Matched	10.69%	0.31	9.90%	0.30	2.09	−66.67%	.56
Acquaintance	Total	53.65%	0.50	60.16%	0.49	10.72		.00
	Matched	57.25%	0.49	56.67%	0.50	0.81	−92.47%	.79
Adult female	Total	20.29%	0.40	19.06%	0.39	2.53		.38
	Matched	21.47%	0.41	20.20%	0.40	2.47	−2.38%	.48
Male child	Total	7.37%	0.26	5.76%	0.23	5.21		.06
	Matched	6.67%	0.25	7.16%	0.26	1.54	−70.43%	.66
Length of stay	Total	36.63	24.50	23.31	21.25	46.33		.00
(months)	Matched	29.90	20.68	29.42	24.44	1.92	−95.86%	.64
Discipline	Total	0.99	1.94	1.46	2.92	16.61		.00
	Matched	1.10	2.02	1.09	1.89	0.34	−97.96%	.94
Supervision	Total	44.60	31.19	33.61	30.07	29.11		.00
(months)	Matched	39.14	28.84	37.88	31.86	3.94	−86.45%	.35
Intensive supervised	Total	33.29%	0.47	17.79%	0.38	28.52		.00
release	Matched	24.71%	0.43	23.33%	0.42	2.51	−91.20%	.47
Supervised release	Total	65.91%	0.47	80.14%	0.40	25.79		.00
	Matched	74.31%	0.44	75.59%	0.43	1.70	−93.39%	.51
Supervised release	Total	0.93	0.27	0.73	0.13	12.63		.00
revocations	Matched	0.85	0.21	0.83	0.17	1.60	−87.37%	.69
Community	Total	7.64%	3.10	1.67%	4.14	21.26		.00
notification	Matched	4.41%	3.20	2.84%	4.12	7.33	−65.51%	.58
Release year	Total	1997.86	1.44	1996.88	1.41	23.09		.00
	Matched	1997.34	1.44	1997.07	1.46	0.86	−96.27%	.09
Risk score	Total	3.77	1.88	3.93	1.83	7.17		.01
	Matched	3.80	1.89	3.84	1.86	1.83	−74.47%	.61

Note: Total treated, $N = 1,493$; total untreated, $N = 1,737$; matched treated, $N = 1,020$; matched untreated, $N = 1,020$.

obtain matches for all the treated offenders. However, using a relatively narrow caliper of .10, we were still able to achieve 1,020 matches, which amounts to 68% of the total number of treated offenders ($N = 1,493$).

Figure 1
Distribution of Propensity Scores by Treatment Assignment

As shown in Table 3, the matching procedure reduced the bias in the propensity score (i.e., probability of entering treatment) by 98%. Again, we present the means for risk score even though it was not used as a predictor in the logistic regression model. In the unmatched sample, more than half of the covariates (9) had bias values greater than 20, and all but three were significantly different at the .05 level. In the matched sample, however, the covariates are balanced to the extent that all bias values are less than 20, and there are no statistically significant differences in covariates between the treated and untreated offenders. The average reduction in bias for the 18 covariates (including risk score) was 81%.

Measures

Dependent variable. Recidivism, the outcome variable, was measured nine different ways in this study. It was first operationalized as (1) rearrest, (2) reconviction, or (3) reincarceration in a MCF for a new offense following an offender's first release from prison. Because it is important to know whether offenders recidivate with a sex offense, recidivism was further distinguished by the type of reoffense: (1) sex

offense, (2) violent offense (including sex offenses), and (3) any offense. Sex offense was defined here as a first- to fifth-degree criminal sexual conduct (CSC) offense. According to Minnesota statutes, CSC first to fourth degree are felony-level offenses, and CSC fifth-degree is a gross misdemeanor offense. In addition to sex crimes, violent offenses included homicide, assault, robbery, and kidnapping.

Arrest, conviction, and incarceration data were collected on offenders through December 31, 2006. The minimum follow-up period, then, was 3 years, whereas the maximum was 17 years. Data on arrests (misdemeanor, gross misdemeanor, and felony) and convictions (misdemeanor, gross misdemeanor, and felony) were obtained electronically from the Minnesota Bureau of Criminal Apprehension, whereas incarceration data were derived from the MNDOC's Correctional Operation Management System (COMS) database. Consequently, a limitation with these data is that they measure only arrests, convictions, or incarcerations that took place in the state of Minnesota. Moreover, as with any recidivism study, official criminal history data will likely underestimate the actual extent to which the sex offenders examined here recidivated.

An arrest, conviction, and/or incarceration was considered a recidivism event only if it pertained to an offense that had taken place following release. There were a handful of offenders who returned to prison for a "new" sex offense that had been committed prior to the beginning of their previous prison term, for example, an offender who was incarcerated from 1997 to 2000 (the beginning of the at-risk period) returns to prison in 2002 for an offense committed in 1995. In these instances, the offenses were not considered recidivism events, but the time that offenders served in prison was deducted from their at-risk period.

Treatment variables. In the statistical analyses presented later, recidivism is the dependent variable. Given that the central purpose of this study is to determine whether sex offender treatment has an impact on recidivism, treatment is the principal variable of interest. In an effort to acquire a more refined understanding of its potential effect on recidivism, we used two separate treatment measures.

The first treatment variable compared offenders who entered sex offender treatment with a comparison group of similar offenders who did not. As such, treatment was measured as "1" for treatment participants and as "0" for nonparticipants. The second treatment variable measured the impact of treatment outcome on reoffending. To this end, we created three dichotomous dummy variables: completion/ successfully participated until the time of release (1 = *completion/successful participation*; 0 = *treatment dropout or nonparticipants*), terminated from treatment or voluntarily quit (1 = *treatment terminations/quits*; 0 = *other*), and nonparticipants (1 = *comparison group*; 0 = *treatment participants*).

Control variables. The control, or independent, variables included in the statistical models were those that were not only available in the COMS database but also

might theoretically have an impact on whether an offender recidivates. Prior research indicates that sex offender recidivism is predicted by factors such as prior sexual criminal history, victim characteristics, the intensity and length of postrelease supervision, and broad community notification (Duwe & Donnay, 2008; Hanson & Morton-Bourgon, 2004; Minnesota Department of Corrections, 2007). To control for potential rival causal factors, it was necessary to include variables such as these in the statistical analyses. The following lists these variables, which include pretreatment and posttreatment measures, and describes how they were created. The univariate relationships between these variables and the three types of recidivism are presented in the appendix.

- *Offender race*: dichotomized as White (0) or minority (1).
- *Age at release*: the age of the offender in years at the time of release based on the date of birth and release date.
- *Prior felony convictions*: offenders who had at least one prior felony conviction (excluding the instant offense) were given a value of 1, whereas those without a prior felony conviction were assigned a value of 0.
- *Prior sex crime convictions*: offenders who had at least one prior sex crime conviction (excluding the instant offense) were given a value of 1, whereas those without a prior sex crime conviction were assigned a value of 0.
- *Victim–offender relationship*: three dichotomous dummy variables were created to measure the offender's relationship to the victim for the instant sex offense, that is, the crime for which the offender was incarcerated. The three variables were stranger victims (1 = *stranger victim*; 0 = *known or nonstranger victim*), acquaintance victims (1 = *acquaintance victim*; 0 = *nonacquaintance victim*), and family member victims (1 = *family member victim*; 0 = *non–family member victim*). The family member victim variable, which is a proxy for incest offenders, serves as the reference in the statistical analyses.
- *Male child victims*: dichotomized as either male child victims (1) or nonmale child victims (0), this variable measures whether offenders victimized a male less than the age of 13 in their instant offense.
- *Adult female victims*: dichotomized as either adult female victims (1) or nonadult female victims (0), this variable quantifies whether offenders victimized women more than the age of 17 in their instant offense. Accordingly, this variable is a proxy for adult rapists.
- *Metro area*: a rough proxy of urban and rural Minnesota, this variable measures an offender's county of commitment, dichotomizing it into either metro area (1) or Greater Minnesota (0). The seven metro area (i.e., Minneapolis, St. Paul, and surrounding suburbs) counties are Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington. The remaining 80 counties were coded as non–metro area or Greater Minnesota counties.
- *Recent disciplinary history*: this variable measures the number of formal disciplinary convictions that an offender received in the final 12 months prior to his initial release from prison. Because sex offenders often serve relatively long sentences, disciplinary convictions at the end of their term of imprisonment may be a more

valid predictor of postrelease behavior than the total number of convictions throughout the full prison term.

- *Risk score*: because formal risk assessment data were not available for the full 14-year period over which offenders were released, we followed the approach developed by Hanson, Broom, and Stephenson (2004) and created a recidivism risk score using the data available from the 10 preceding control variables. Offenders received a value of "1" if they were less than the age of 30 at the time of release, had at least one institutional discipline conviction in the 12 months prior to release, or had a value of "1" for the remaining eight controls (minority race, prior felony conviction, prior sex crime conviction, stranger victims, acquaintance victims, male child victims, adult female victims, and metro area). Thus, the maximum total score was 10, whereas the minimum score was 0.
- *Length of stay (LOS)*: the number of months between prison admission and release dates.
- *Length of postrelease supervision*: the number of months between an offender's first release date and the end of postrelease supervision, that is, the sentence expiration or conditional release date, the greater of the two.
- *Type of postrelease supervision*: three dichotomous dummy variables were created to measure the level of postrelease supervision to which offenders were released. The three variables were intensive supervised release (ISR; 1 = *ISR*; 0 = *non-ISR*), supervised release (SR; 1 = *SR*; 0 = *non-SR*), and discharge (1 = *discharge or no supervision*; 0 = *released to supervision*). Discharge is the variable that serves as the reference in the statistical analyses.
- *Supervised release revocations*: the number of times during an offender's sex crime sentence when he returned to prison as a supervised release violator for a technical violation.
- *Broad community notification*: dichotomized as either (1) broad community notification or (0) no broad community notification, this variable measures whether offenders were given a Level III risk level assignment prior to their release from prison and, thus, were subjected to broad community notification.
- *Release year*: measuring the year in which offenders were first released from prison for the instant sex offense, this variable is included to control for any unobserved differences between the 14 different release year cohorts from 1990 to 2003.

Analysis

In analyzing recidivism, survival analysis models are preferable in that they use time-dependent data, which are important in determining not only whether offenders recidivate but also when they recidivate. As a result, the statistical technique we used was a Cox regression model, which uses both "status" and "time" variables in estimating the impact of the independent variables on recidivism. For the analyses presented here, the "status" variable was one of the recidivism variables mentioned above, for example, sex crime rearrest, violent crime rearrest. The "time" variable, on the other hand, measured the amount of time (in days) from the date of release until the date of first rearrest, reconviction, reincarceration, or December 31, 2006, for those who did not recidivate.

To accurately measure the total amount of time an offender was actually at risk to reoffend (i.e., "street time"), it was necessary to account for instances in which an offender was not at risk to recidivate following release from prison. Failure to do so would bias the findings by artificially increasing the lengths of offenders' at-risk periods. Accordingly, the time offenders spent in prison as supervised release violators was subtracted from their total at-risk period as long as it (1) preceded a reincarceration for a new offense or (2) occurred prior to January 1, 2007 (the end of the follow-up period) for those who did not recidivate. In addition, when recidivism was defined as a sex reoffense, time spent in prison was deducted for offenders reincarcerated for either a violent or a nonsex reoffense.

Because civilly committed offenders are incapacitated in a mental health institution, it was necessary to account for those who were civilly committed between September 1991—when Minnesota courts reenacted the civil commitment statute on released sex offenders—and December 31, 2006. Of the 3,533 sex offenders released from Minnesota prisons between 1990 and 2003, 93 were excluded because they were later civilly committed without ever spending any time in the community. Of these, 70 entered prison-based treatment, with 32 dropping out and the remaining 38 completing or participating until release. As expected, these offenders had a higher average risk score (4.82) than the other 3,440 offenders (3.89), which suggests that they had a greater recidivism risk than the sex offender population in general.

We included in the study 54 offenders who had spent time in the community but had later been civilly committed following a return to prison for either a supervised release violation or a new crime. Offenders who returned to prison for a supervised release revocation were "right censored" at the time of their civil commitment, that is, their at-risk period ended when they were civilly committed. For offenders who were civilly committed following a reincarceration for a new offense, they were right censored at the time of their commitment if the offense type was different from the type of recidivism being measured (e.g., sexual or violent). For example, when recidivism was measured as a violent reoffense, offenders were right censored at the time of their civil commitment following a return to prison for either a new sex or a nonsex crime.

We estimated Cox regression models for each of the nine recidivism measures for both treatment variables (participation and outcome). However, given that the reconviction and reincarceration results were substantively similar to those for rearrest for all three reoffense types, we present only the findings for rearrest because it is the most sensitive recidivism measure. Nevertheless, the reconviction and reincarceration results can be obtained from the authors on request. In addition, to determine whether there are certain types of offenders for whom treatment may be more effective, we estimated interaction models for each measure of recidivism. Similar to stepwise regression, we examined all first-order interactions with treatment and removed nonsignificant terms until only the significant interactions (at the .05 level) remained in the model.

Results

As shown in Table 4, which breaks out recidivism rates by treatment participation and outcome, treated offenders had lower reoffense rates than untreated offenders for each of the three types of recidivism—sexual, violent, and general. Not surprisingly, the best recidivism outcomes were found for offenders who completed treatment or successfully participated until their release. These results suggest that the risk of recidivism may be significantly lowered by participating in prison-based treatment, especially for those who complete treatment or successfully participate until release. It is possible, however, that the observed recidivism differences between treated and untreated offenders as well as between treatment completers and dropouts are due to other factors such as prior criminal history, discipline history, or postrelease supervision. To statistically control for the impact of these other factors on reoffending, we estimated Cox regression models for each measure of recidivism across both treatment variables (participation and outcome).

The Impact of Treatment on Sex Offender Recidivism

For each measure of recidivism, we initially ran two separate Cox regression models to estimate the effects of prison-based treatment. The first model, risk score, contained the 10-factor risk score measure along with the institutional and postrelease controls. The individual predictor model, on the other hand, was similar to the risk score model except that it showed the unique effects of the 10 predictors used to calculate the risk score. Because the results from the individual predictor models were similar to those from the risk score models for all three types of recidivism, only the findings from the risk score models are presented here. Results from the individual predictor models can be obtained, however, from the authors on request.

Sexual recidivism. The results shown in Table 5 indicate that, controlling for other factors, prison-based treatment significantly reduced the hazard ratio for a new sex offense rearrest, decreasing it by 27%. That is, sex offenders who participated in treatment recidivated less often and more slowly than untreated offenders; as a result, treated sex offenders survived longer in the community without committing a new sex offense (see Figure 2). In the individual predictor model, the hazard ratio was 28% lower for treatment participants.

Although not shown in Table 5, we also estimated Cox regression models that analyzed the impact of treatment outcome on sexual recidivism. Compared with the untreated offenders, the effect of dropping out of treatment—either quitting or being terminated—was in the negative direction and did not have a significant effect on sexual recidivism. Completing treatment, however, did significantly decrease the risk (hazard) relative to not receiving treatment, reducing it by 33% in the risk score model and 34% in the individual predictor model.

Table 4
Three-Year and Total Recidivism Rates
by Treatment Participation and Outcome

Recidivism	Treatment Completers	Treatment Dropouts	Treatment Participants	Treatment Nonparticipants
Sexual rearrest				
Three years	7.1% (51)	10.6% (32)	8.1% (83)	11.6% (118)
Total	13.4% (96)	16.2% (49)	14.2% (145)	19.5% (199)
Violent rearrest				
Three years	13.4% (96)	16.9% (51)	14.4% (147)	19.3% (197)
Total	29.0% (208)	35.1% (106)	30.8% (314)	34.1% (348)
General rearrest				
Three years	29.1% (209)	33.1% (100)	30.3% (309)	38.5% (393)
Total	55.4% (398)	59.3% (179)	56.6% (577)	58.1% (593)
N	718	302	1,020	1,020

Table 5
Cox Regression Model: Time to First Sex Offense Rearrest

Variables	Coefficient	SE	Hazard Ratio
Prison-based treatment	−0.317	0.110	0.729**
Risk score	0.135	0.030	1.145**
Length of stay (months)	−0.010	3.0E-3	0.990**
Supervision length (months)	−3.1E-4	2.3E-3	1.000
Intensive supervised release	−1.041	0.358	0.353**
Supervised release	−1.484	0.327	0.227**
Supervised release revocations	−0.066	0.049	0.936
Community notification	−1.242	0.589	0.289*
Release year	−0.093	0.021	0.911**
N	2,040		

$*p < .05.$ $**p < .01.$

We tested for interactions between the controls and the two treatment variables (participation and outcome) in both the risk score and individual predictor models, but none were statistically significant. The results from all four models, however, showed that longer lengths of stay in prison, postrelease supervision (ISR or supervised release), broad community notification, and release year were associated with a reduced risk of rearrest. The findings from both risk score models (treatment participation and outcome) revealed that risk score was a significant predictor of sexual recidivism; in the treatment participation model, a one unit increase in risk score increased the hazard ratio for sexual recidivism by 15%. In the two individual predictor models, minority offenders, younger offenders, and prior sex crime convictions were significantly associated with an increased risk of sexual recidivism.

Figure 2
Survival Curves for Sexual Rearrest
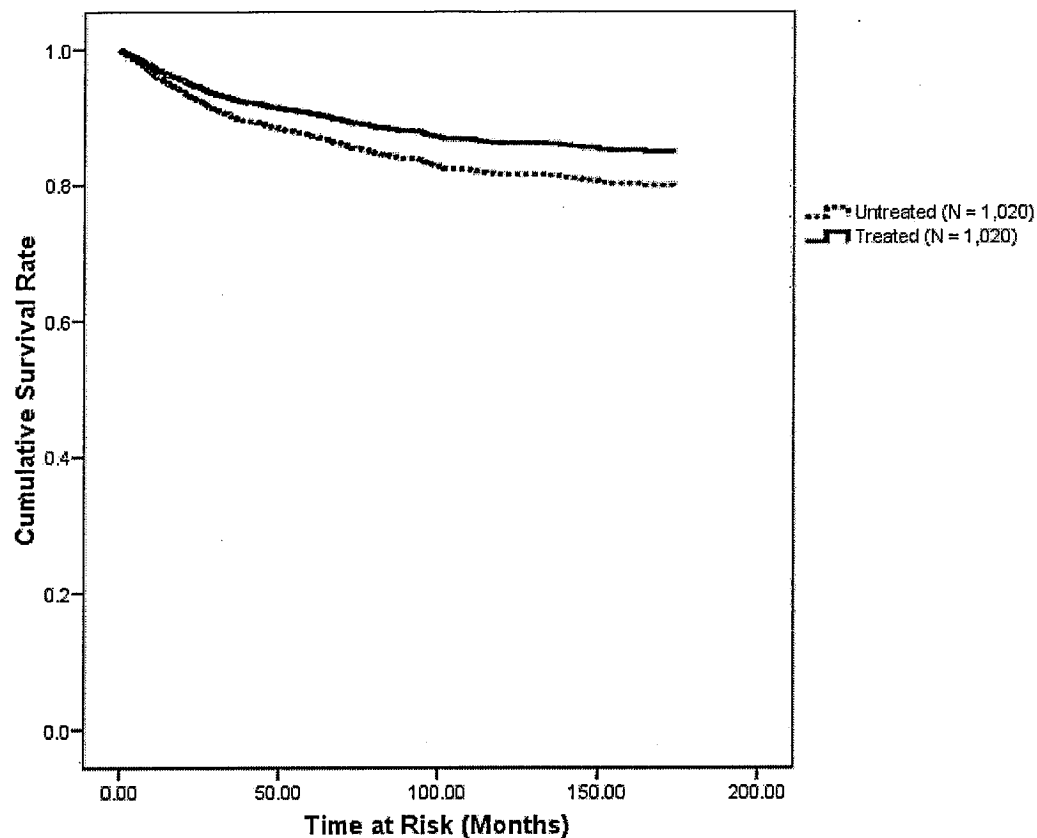
Violent recidivism. The results in Table 6 show that treatment had a statistically significant impact on violent offense recidivism. The hazard ratio for a violent rearrest was 18% lower for treated sex offenders in the risk score model and 19% lower in the individual predictor model (also see Figure 3). Compared with the untreated offenders, completing treatment reduced the risk (hazard) by 23% in the risk score model and by 24% in individual predictor model, whereas the effect of dropping out of treatment was in the negative direction and was not significant in either model. Similar to the sexual recidivism results, we did not find any statistically significant interactions between any of the controls and either treatment variable.

The findings from all four models (treatment participation/risk score, treatment participation/individual predictor, treatment outcome/risk score, and treatment outcome/individual predictor) suggested that longer postrelease supervision periods, postrelease supervision (ISR and supervised release), and release year were significantly associated with a reduced risk of rearrest for a violent offense. Supervised release

Table 6
Cox Regression Model: Time to First Violent Offense Rearrest

Variables	Coefficient	SE	Hazard Ratio
Prison-based treatment	−0.194	0.079	0.824*
Risk score	0.157	0.022	1.170**
Length of stay (months)	−3.5E-3	2.0E-3	0.996
Supervision length (months)	−4.1E-3	1.7E-3	0.996*
Intensive supervised release	−1.379	0.301	0.252**
Supervised release	−1.426	0.284	0.240**
Supervised release revocations	0.169	0.029	1.184**
Community notification	−0.531	0.282	0.588
Release year	−0.039	0.015	0.962**
N	2,040		

$*p < .05.$ $**p < .01.$

Table 7
Cox Regression Model: Time to First Rearrest for Any Offense

Variables	Coefficient	SE	Hazard Ratio
Prison-based treatment	−0.123	0.059	0.884*
Risk score	0.198	0.016	1.219**
Length of stay (months)	−5.2E-3	1.5E-3	0.995**
Supervision length (months)	−3.9E-3	1.3E-3	0.996**
Intensive supervised release	−0.675	0.291	0.501*
Supervised release	−0.536	0.281	0.585
Supervised release revocations	0.083	0.026	1.089*
Community notification	−0.870	0.213	0.419**
Release year	0.026	0.011	1.027*
N	2,040		

$*p < .05.$ $**p < .01.$

revocations, on the other hand, significantly increased the risk in all four models. Risk score was a significant predictor of violent recidivism in both treatment variable models, whereas minority offenders, younger offenders, and prior felony convictions significantly increased the hazard ratio for rearrest in the individual predictor models.

General recidivism. As shown in Table 7, participating in treatment had a statistically significant effect on general recidivism, reducing the hazard ratio for rearrest for any offense by 12% (also see Figure 4). In the individual predictor model, the hazard ratio was 14% lower for treated offenders. The treatment outcome results suggest that, compared with untreated offenders, completing treatment

Figure 3
Survival Curves for Violent Rearrest
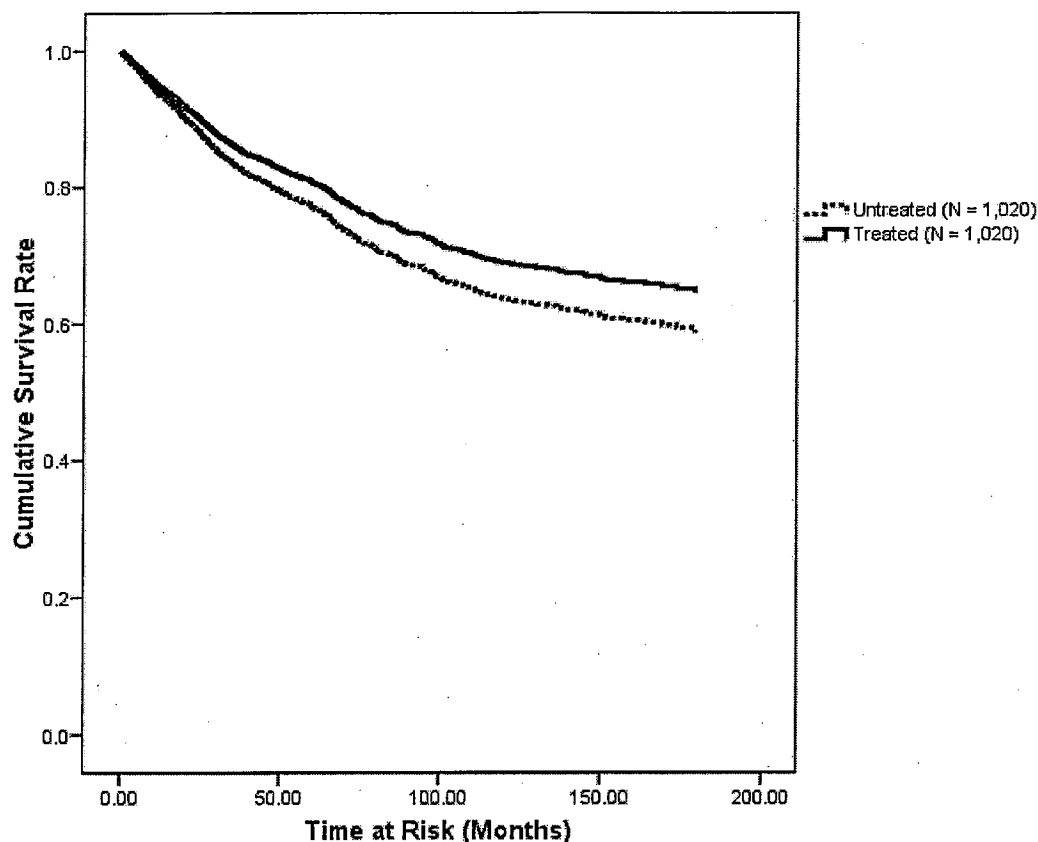
significantly decreased the risk (hazard) for any offense by 15% in the risk score model and by 17% in the individual predictor model. The effect of dropping out of treatment, relative to not participating in treatment, was in the negative direction and was not statistically significant in both models. Once again, we did not find any statistically significant interactions between the controls and either treatment variable.

Longer periods of postrelease supervision, longer lengths of stay in prison, ISR, and broad community notification significantly decreased the hazard ratio for rearrest in all four models, whereas release year and supervised release revocations were significantly and positively associated with recidivism risk. Risk score significantly increased the risk of rearrest in both treatment models. In the individual predictor models, minority offenders, younger offenders, prior felony convictions, a history of victimizing acquaintances, and recent institutional disciplinary convictions

Figure 4
Survival Curves for General Rearrest
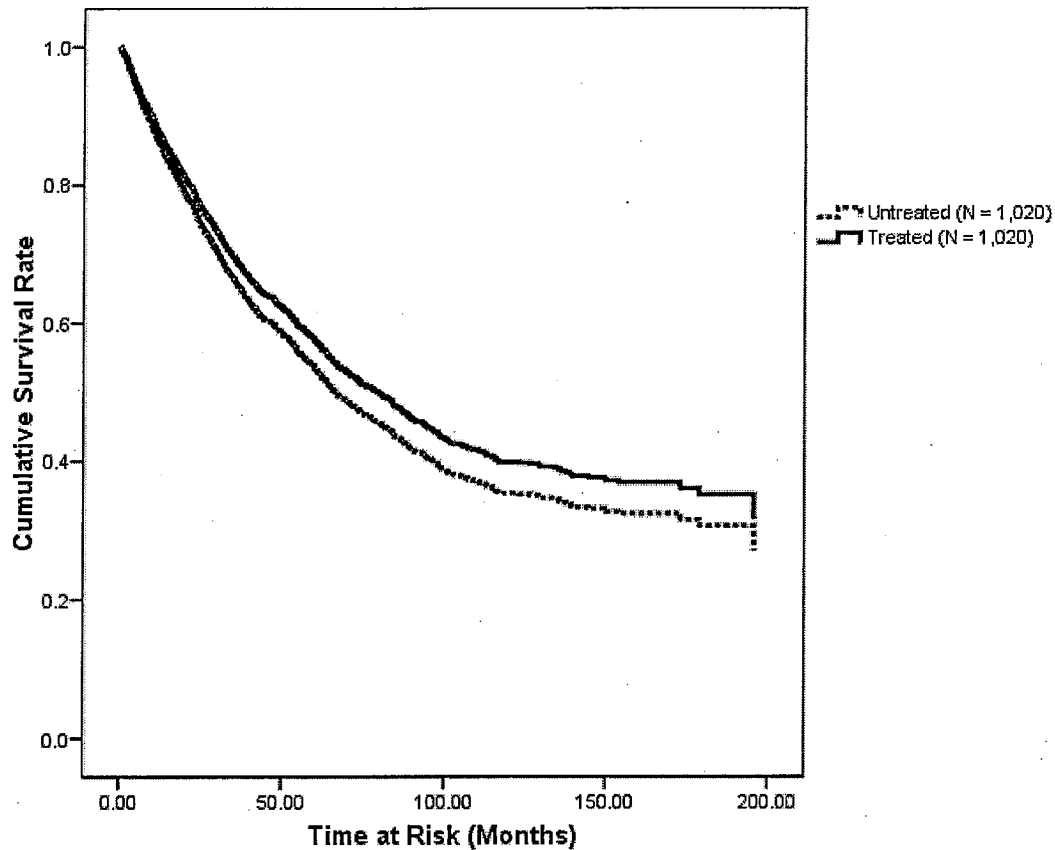
significantly increased the risk of general recidivism. In contrast, prior sex crime convictions and a history of victimizing male children significantly decreased the risk.

Sensitivity Analyses

Intent to treat. The results presented above suggest that prison-based treatment in Minnesota significantly reduces sexual, violent, and general recidivism. But in using PSM to identify the untreated offenders most likely to refuse treatment, it is possible that we may have underestimated the number of offenders who would have refused to enter treatment had it been offered to them. For example, treatment refusers (105) accounted for 6.6% of the offenders (1,598) who were offered treatment. Yet the 105 offenders matched to the treatment refusers made up 5.7% of those not offered treatment (1,842). If the rate of refusal was the same among the 1,842 not offered

treatment, then 121 offenders (6.6% of 1,842) would have refused a treatment offer. The results from the preceding analyses could be biased, then, to the extent that 16 additional offenders needed to be removed from the comparison group pool but were not.

To address potential treatment refuser bias, we conducted intent-to-treat (ITT) analyses based on whether offenders were offered treatment. ITT analysis does not measure the effectiveness of treatment administered insofar as treatment refusers are considered to be "treated" offenders. It can be used, however, to test whether our results are robust to possible treatment refuser bias.

We began the ITT analyses by using PSM to individually match offenders not offered treatment with those who received a treatment offer. We estimated a logistic regression model in which the dependent variable was a treatment offer (i.e., the 1,598 offenders offered treatment were assigned a value of "1," whereas the 1,842 untreated offenders not offered treatment received a value of "0"). The predictors were the 17 control variables (excluding risk score) described earlier. After obtaining propensity scores on the 3,440 offenders, we used the greedy matching procedure to individually match offenders from both groups. Using a caliper of .10, we were able to obtain a sample of 2,224 in which all of the covariates were balanced. The number of matches (1,112) accounted for 70% of the total number of offenders offered treatment ($N = 1,598$). Moreover, of the 1,112 matched pairs, 85 were treatment refusers (81% of the 105 refusers).

Using Cox regression, we estimated risk score models for sexual, violent, and general recidivism. The results from these models, which can be obtained from the authors, showed that the hazard ratios for offenders offered treatment were significantly lower for all three types of recidivism. In particular, the hazard ratios were 19% lower for sexual recidivism ($B = -.209$; $SE = .101$), 14% lower for violent recidivism ($B = -.156$; $SE = .073$), and 16% lower for general recidivism ($B = -.175$; $SE = .056$). Compared with the recidivism analyses reported in the previous section, the hazard ratios were smaller for sexual and violent recidivism but were slightly larger for general recidivism.

Rosenbaum bounds. As indicated by the results from the ITT analyses, the treatment effects were robust against possible treatment refuser bias. Yet, given that PSM controls only for bias among the observed covariates, the possibility exists that unobserved selection bias may account for the significant treatment effects. Hidden bias can occur when two offenders with the same observed covariates have different chances of receiving treatment due to an unobserved covariate. If this unobserved covariate is related to the outcome (recidivism) affected by treatment, then the failure to account for this hidden bias can alter conclusions drawn about the effects of treatment.

We tested the sensitivity of our results to hidden bias by using a method developed by Rosenbaum (2002) that calculates a bound on how large an effect an unobserved covariate would need to have on the treatment selection process to reverse

inferences drawn about the effects of treatment. The Rosenbaum bounds sensitivity analysis produces a test statistic, gamma, that measures the threshold at which an unobserved covariate would cause the estimated treatment effect to no longer be statistically significant (i.e., $p > .05$). More specifically, the closer the gamma value is to 1, the stronger the possibility that the effect can be explained away by an unobserved covariate. Therefore, an estimated treatment effect with a gamma value of, say, 1.5 would be more sensitive to hidden bias than an effect with a gamma value of 2.0.

It is important to emphasize, however, that the Rosenbaum bounds method is limited in two important ways. First, the sensitivity analysis does not indicate whether unobserved bias exists. Rather, it simply identifies how large the hidden bias would need to be to nullify the estimated treatment effect. Second, as DiPrete and Gangl (2004) point out, the Rosenbaum bounds method is a "worst-case" scenario to the extent that it assumes the hypothetical unobserved covariate is an almost perfect predictor of the outcome variable (recidivism).

The results from the sensitivity analyses reveal that the estimated treatment effects are not especially robust to hidden bias. With a gamma value of 1.02, the general recidivism findings are the most sensitive to the possibility of hidden bias, followed by violent recidivism (gamma = 1.09) and sexual recidivism (gamma = 1.15). These results suggest that if an unobserved covariate that almost perfectly predicted general recidivism differed between matched pairs of treated and untreated offenders by a factor of 1.02 or more, it would be sufficient to undermine the conclusions regarding the treatment effect. To put this statistic in perspective, LOS would be a hidden bias equivalent in that, as shown earlier in Table 2, it had a comparable impact on the treatment selection process ($b = .02$). Therefore, if an unobserved covariate existed that perfectly predicted general recidivism and had an impact on the treatment selection process similar to LOS, it would be sufficient to invalidate the treatment effect for general recidivism. Furthermore, most of the significant predictors of treatment selection shown earlier in Table 2 had effect sizes ($b > .15$) that exceeded the gamma value for sexual recidivism (1.15), which was the least sensitive to possible hidden bias. Still, it is worth reiterating, however, that the Rosenbaum bounds method is a "worst-case" scenario. Although existing research has identified a number of factors that are significantly associated with sex offender recidivism, none have yet to be shown to be a nearly perfect predictor of reoffending, which is what the Rosenbaum bounds approach assumes.

Conclusion

The results from this study suggest that prison-based treatment in Minnesota produces a significant, albeit relatively modest, reduction in sex offender recidivism. Indeed, entering treatment lowered the risk of rearrest for a new offense by 12% for

general recidivism, 18% for violent recidivism, and 27% for sexual recidivism. The average sexual recidivism rate was 27% lower for treated offenders (14.2%) than for untreated offenders (19.5%), which is similar to the reduction reported by Hanson et al. (2002) but lower than that (37%) reported by Lösel and Schmucker (2005) in their meta-analyses of the treatment evaluation literature. Moreover, the effect size for sexual recidivism ($d = .21$), which translates to an odds ratio of 1.46 (Lösel & Schmucker, 2005; Sánchez-Meca, Marín-Martínez, & Chacón-Moscoso, 2003), falls within the lower end of the range ($d = .12-.47$) observed in previous meta-analytic reviews.

Dropping out of treatment did not significantly increase the risk of recidivism, although completing treatment lowered it for sexual rearrest by 33%, violent rearrest by 23%, and any arrest by 15%. In addition, we did not find treatment to be significantly more, or less, effective for certain types of sex offenders. The absence of any significant interactions is important in its own right, however, for it suggests that treatment is similarly effective not only for adult rapists and child molesters but also for incest offenders and those who victimize acquaintances or strangers.

Although the findings reported here support the notion that prison-based sex offender treatment is moderately effective in Minnesota, there are several limitations worth reiterating. First, due to a lack of data, this study did not account for the potential impact that community-based treatment may have had on recidivism. Prior research has shown that community-based treatment significantly lowers the extent to which sex offenders recidivate, particularly with regard to sex offenses (Aytes, Olsen, Zakrajsek, Murray, & Ireson, 2001; Marshall & Barbaree, 1988; Marshall, Eccles, & Barbaree, 1991; McGrath, Hoke, & Vojtisek, 1998). It is possible, then, that the sexual recidivism differences between treated and untreated offenders may reflect variations in the extent to which each group participated in community-based treatment, that is, more prison-treated offenders may have participated in community-based treatment than untreated offenders.

Recall, however, that sex offenders who do not receive treatment while incarcerated are directed, as part of their postrelease supervision conditions, to enter treatment while in the community. As a result, it is unlikely that prison-treated offenders were significantly more likely to participate in community-based treatment than offenders who were not treated in prison. If anything, sex offenders who were untreated in prison likely had a higher rate of participation in community-based treatment than prison-treated offenders. Accordingly, if community-based treatment is as effective in Minnesota as suggested by prior research, it may have moderated the observed effect for prison-based treatment. Therefore, the actual effect of prison-based treatment may be stronger than what was reported in this study.

Second, because this study did not use a randomized experimental design, some may argue that it does not provide an adequate assessment of the effectiveness of prison-based treatment—in Minnesota or in general. However, random assignment does not guarantee equivalence between treated and untreated offenders (Marques

et al., 2005). Moreover, as Hanson, Broom, and Stephenson (2004) point out, no single study—regardless of how rigorous the design—is sufficient to determine whether treatment works. Instead, Hanson and colleagues argue that advances in the understanding of sex offender treatment will be made when individual studies improve and the cumulative results from these studies are meaningfully integrated through meta-analyses. This study thus contributes to the advancement of the sex offender treatment literature by not only examining a relatively large number of sex offenders but also by using multiple treatment and outcome measures, a lengthy follow-up period, a matching technique that controlled for observable selection bias, and sensitivity analyses that addressed treatment refuser and hidden selection bias.

Appendix
Univariate Relationships Between Control Variables and Recidivism

Control Variables Dichotomous Variables	Sexual Rearrest Odds Ratios	Violent Rearrest Odds Ratios	Any Rearrest Odds Ratios
Minority	1.002	1.658**	1.804**
Metro	1.314**	1.359**	1.346**
Prior sex crimes	1.363**	0.903	0.746**
Prior felony	1.161	1.451**	1.997**
Stranger	1.369*	1.498**	1.246
Acquaintance	1.152	1.311**	1.676**
Adult female	1.092	1.387**	1.427**
Male child	1.272	0.762	0.608**
Intensive supervised release	0.788*	0.660**	0.640**
Supervised release	1.082	1.331**	1.508**
Community notification	0.608*	0.632*	0.476**
Ordinal/Interval Variables	AUC	AUC	AUC
Age at release (years)	0.360**	0.404**	0.455**
Length of stay (months)	0.467*	0.461**	0.425**
Discipline	0.543**	0.583**	0.595**
Supervision (months)	0.443**	0.415**	0.438**
Supervised release revocations	0.522	0.579**	0.562**
Release year	0.377**	0.371**	0.400**
Risk score	0.564**	0.622**	0.649**
N	3,440	3,440	3,440

Note: AUC = area under the curve.
*$p < .05$. **$p < .01$

References

Alexander, M. A. (1999). Sexual offender treatment efficacy revisited. *Sexual Abuse: A Journal of Research and Treatment, 11*, 101-116.

Aytes, K. E., Olsen, S. S., Zakrajsek, T., Murray, P., & Ireson, R. (2001). Cognitive/behavioral treatment for sexual offenders: An examination of recidivism. *Sexual Abuse: A Journal of Research and Treatment, 13*, 223-231.

Caldwell, M., Skeem, J., Salekin, R., & Van Rybroek, G. (2006). Treatment response of adolescent offenders with psychopathy features. *Criminal Justice and Behavior, 33*, 571-596.

DiPrete, T. A., & Gangl, M. (2004). Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological Methodology, 34*, 271-310.

Duwe, G., & Donnay, W. (2008). The impact of Megan's law on sex offender recidivism: The Minnesota experience. *Criminology, 46*, 411-446.

Furby, L., Weinrott, M.R., & Blackshaw, L. (1989). Sex offender recidivism: A review. *Psychological Bulletin, 105*, 3-30.

Gallagher, C. A., Wilson, D. B., Hirschfield, P., Coggeshall, M. B., & MacKenzie, D. L. (1999). A quantitative review of the effects of sex offender treatment on sexual reoffending. *Corrections Management Quarterly, 3*, 19-29.

Hall, G. C. N. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology, 63*, 802-809.

Hanson, R. K., Broom, I. R., & Stephenson, M. (2004). Evaluating community sex offender treatment programs: A 12-year follow-up of 724 offenders. *Canadian Journal of Behavioural Science, 37*, 87-96.

Hanson, R. K., Gordon, A., Harris, A. J. R., Marques, J. K., Murphy, W. D., & Quinsey, V.L. (2002). First report on the collaborative outcome data project on the effectiveness of psychological treatment for sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 14*, 169-194.

Hanson, R. K., & Morton-Bourgon, K. (2004). Predictors of Sexual Recidivism: An updated meta-analysis (User Rep. No. 2004-02). Ottawa, Ontario, Canada: Department of the Solicitor General of Canada.

Harkins, L., & Beech, A. (2006). Measurement of the effectiveness of sex offender treatment. *Aggression and Violent Behavior, 12*, 36-44.

Jones, N., Pelissier, B., & Klein-Saffran, J. (2006). Predicting sex offender treatment entry among individuals convicted of sexual offense crimes. *Sexual Abuse: A Journal of Research and Treatment, 18*, 83-98.

Kenworthy, T., Adams, C. E., Brooks-Gordon, B., & Fenton, M. (2004). Psychological interventions for those who have sexually offended or are at risk of offending. *Cochrane Database of Systematic Reviews*, Issue 3: CD004858.

Lösel, F., & Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology, 1*, 117-146.

Marques, J. K. (1999). How to answer the question, "Does sex offender treatment work?" *Journal of Interpersonal Violence, 14*, 437-451.

Marques, J. K., Day, D. M., Nelson, C., & West, M. A. (1994). Effects of cognitive-behavioral treatment on sex offender recidivism: Results of a longitudinal study. *Criminal Justice and Behavior, 21*, 28-54.

Marques, J. K., Wiederanders, M., Day, D. M., Nelson, C., & von Ommeren, A. (2005). Effects of a relapse prevention program on sexual recidivism: Final results from California's Sex Offender Treatment and Evaluation Project (SOTEP). *Sexual Abuse: A Journal of Research and Treatment, 17*, 79-107.

Marshall, W., & Barbaree, H. (1988). The long-term evaluation of a behavioral treatment program for child molesters. *Behavior Research and Therapy, 26*, 499-511.

Marshall, W., Eccles, A., & Barbaree, H. (1991). The treatment of exhibitionists: A focus on sexual deviance versus cognitive and relationship features. *Behavior Research and Therapy, 29*, 129-135.

McGrath, R. J., Hoke, S. E., & Vojtisek, J. E. (1998). Cognitive-behavioral treatment of sex offenders. *Criminal Justice and Behavior, 25*, 203-225.

Minnesota Department of Corrections. (2007). *Sex offender recidivism in Minnesota*. St. Paul, MN: Author.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Lalumiere, M. L. (1993). Assessing treatment efficacy in outcome studies of sex offenders. *Journal of Interpersonal Violence, 8,* 512-523.

Rice, M. E., & Harris, G. T. (2003). The size and sign of treatment effects in sex offender therapy. *Annals of the New York Academy of Sciences, 989,* 428-440.

Rosenbaum, P. (2002). *Observational studies* (2nd ed.). New York: Springer.

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association, 79,* 516-524.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician, 39,* 33-38.

Rubin, D. B. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine, 127,* 757-763.

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods, 8,* 448-467.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Skeem, J. L., Monahan, J., & Mulvey, E. P. (2002). Psychopathy, treatment involvement, and subsequent violence among civil psychiatric patients. *Law and Human Behavior, 26,* 577-603.

First Report of the Collaborative Outcome Data Project on the Effectiveness of Psychological Treatment for Sex Offenders[1]

R. Karl Hanson,[2,8] Arthur Gordon,[3] Andrew J. R. Harris,[2] Janice K. Marques,[4] William Murphy,[5] Vernon L. Quinsey,[6] and Michael C. Seto[7]

This meta-analytic review examined the effectiveness of psychological treatment for sex offenders by summarizing data from 43 studies (combined n = 9,454). Averaged across all studies, the sexual offence recidivism rate was lower for the treatment groups (12.3%) than the comparison groups (16.8%, 38 studies, unweighted average). A similar pattern was found for general recidivism, although the overall rates were predictably higher (treatment 27.9%, comparison 39.2%, 30 studies). Current treatments (cognitive–behavioral, k = 13; systemic, k = 2) were associated with reductions in both sexual recidivism (from 17.4 to 9.9%) and general recidivism (from 51 to 32%). Older forms of treatment (operating prior to 1980) appeared to have little effect. Future directions for improving the quality of sex offender treatment outcome evaluations are discussed.

KEY WORDS: sex offender treatment; evaluation; recidivism; meta-analysis.

[1]The views expressed are those of the authors and do not necessarily reflect the positions of the Association for the Treatment of Sexual Abusers, the Department of the Solicitor General of Canada, the Correctional Service of Canada, or the California Department of Mental Health.
[2]Corrections Research, Department of the Solicitor General of Canada, Ottawa, Canada.
[3]Clearwater Treatment Centre, Monroe, Washington.
[4]California Department of Mental Health, Sacramento, California.
[5]University of Tennessee–Memphis, Memphis, Tennessee.
[6]Queen's University, Kingston, Ontario.
[7]Centre for Addiction and Mental Health, Toronto, Ontario.
[8]To whom correspondence should be addressed at Corrections Research, Department of the Solicitor General of Canada, 340 Laurier Avenue West, Ottawa, Canada K1A 0P8; e-mail: hansonk@sgc.gc.ca.

If treatment is to be widely used in the management of sex offenders, then it is important that it works. Despite more than 35 review papers since 1990, and a review of reviews (United States General Accounting Office, 1996), researchers and policy-makers have yet to agree on whether treatment effectively reduces sexual recidivism.

This lack of consensus is rooted in inherent difficulties with the evaluation of sex offender treatment programs. On the surface, the evaluation process appears straightforward: The sexual offence recidivism rate of a treated group of sex offenders should be compared to the recidivism rate of an equivalent group of offenders who did not receive treatment. Researchers who attempt such comparisons are faced, however, with the challenge of low recidivism rates even among untreated offenders (Barbaree, 1997). On average, only 10–15% of sex offenders are detected committing a new sexual offence after 4–5 years (Hanson & Bussière, 1998). The typical treatment program provides service to relatively few offenders. To achieve sufficient statistical power, researchers who initiate new studies have to wait many years before treatment effects can be detected. Those who choose to evaluate existing programs must use post hoc comparison groups, often of questionable comparability.

One way to increase statistical power is to aggregate studies through meta-analysis (e.g., Cooper & Hedges, 1994). By combining the findings of numerous individual studies, the resulting sample size can be sufficient to detect even small effects. To date, three meta-analyses of the treatment outcome literature for sex offenders have been published (Alexander, 1999; Gallagher, Wilson, Hirschfield, Coggeshall, & Mackenzie, 1999; Hall, 1995b).

Alexander (1999) summarized the recidivism rates of separate groups of untreated and treated sex offenders. The treated and untreated sex offenders, in most cases, came from different studies. Consequently, it was difficult to tell whether any observed differences could be attributed to treatment effects or to differences in follow-up periods, offender samples, recidivism criteria, or other design features. Although she concluded there was evidence for an overall treatment effect, Alexander's results contained some anomalies. For example, she found lower recidivism rates for treated versus untreated child molesters, and for treated versus untreated rapists, but there was no significant treatment effect for combined samples of rapists and child molesters. Critics could argue there was too much method variance across studies to allow for clear conclusions.

Hall's meta-analysis attempted to control for this method variance by including only studies that included a comparison group (Hall, 1995b). Rather than summarizing overall recidivism rates, Hall defined the difference in recidivism rates between each treatment and comparison group as the outcome criterion. Hall only examined studies ($k = 12$) that appeared after Furby, Weinrott and Blackshaw's discouraging narrative review of the early (largely pre-1980) treatment outcome literature (Furby, Weinrott, & Blackshaw, 1989). This meta-analysis found a small, but significant, overall treatment effect ($r = .12$), and concluded that medical

treatments and comprehensive cognitive–behavioral treatments were both superior to purely behavioral treatments.

A major criticism of Hall's meta-analysis is that the strongest treatment effects came from comparisons between treatment completers and dropouts (Hall, 1995b). Such comparisons are difficult to interpret because those who drop out of treatment are likely to have characteristics related to recidivism risk, such as youth, impulsivity, and antisocial personality (Wierzbicki & Pekarik, 1993). When the dropout studies were removed from Hall's meta-analysis, the treatment effect was no longer significant (Hall, 1995b; Harris, Rice, & Quinsey, 1998).

Of the three meta-analyses, Gallagher et al.'s is the most comprehensive and technically sophisticated (Gallagher et al., 1999). Gallagher et al. considered 25 comparison-group studies examining psychological or hormonal treatments. Like Hall (1995b), they concluded that there was a significant treatment effect for cognitive–behavioral treatments. Unlike Hall, they found insufficient evidence to support medical/hormonal treatments. The apparent effectiveness of medical/hormonal treatments in Hall's review could be attributed to a single study of physical castration in Germany (Wille & Beier, 1989).

Although Gallagher et al. (1999) made some effort to restrict their analysis to well-controlled studies, they nevertheless included a number of studies (6 out of 25) in which bias could be expected (e.g., comparing completers to treatment dropouts). As well, they included early studies of programs that showed greater treatment effects than those revealed in subsequent analyses. For example, the more recent treatment effects reported by California's experimental program (Marques & Day, 1998) were weaker than the effect reported in the earlier study (Marques, Day, Nelson, & West, 1994) that Gallagher et al.'s analysis included. Another study that Gallagher et al. consider to provide "strong evidence" for cognitive–behavioral treatment (Nicholaichuk, Gordon, Andre, & Gu, 1995) contained a subtle methodological bias that may have inflated the treatment effect (Hanson & Nicholaichuk, 2000). Because the comparison group were released earlier than the treatment group, the records of the nonrecidivists were more likely to go missing from the comparison group than from the treatment group.

Gallagher et al.'s review has not resolved the controversy concerning treatment effectiveness and it is unlikely that the controversy will be resolved in the near future (Gallagher et al., 1999). All studies, no matter how well conceived, are open to alternate interpretations; experts can legitimately disagree as to the information value of studies containing more or less serious threats to validity (Hanson, 1997). In this context, the Collaborative Outcome Data Project Committee[9] was

[9]The committee members (listed as the coauthors of this paper) are researchers committed to the empirical evaluation of sex offender treatment programs who met through the Association for the Treatment of Sexual Abusers (ATSA). Although ATSA has provided administrative support for this project (e.g., providing meeting rooms at its annual conferences, advertising the project), the goals of the collaborative project were scientific, and the committee's reports should not be considered products of ATSA or any other organization.

formed in 1997, with the goals of organizing the existing outcome literature and encouraging new evaluation projects to be conducted in a manner that contributes to scientific knowledge.

The committee's first challenge was distinguishing between research designs on the basis of potential threats to validity. A detailed coding manual was developed that included over 100 variables describing the treatments and another 100 variables describing the research designs. Available studies, however, rarely collected and reported such detailed information. Consequently, the analyses reported in the current report were based on broad classifications. These broad classifications obscure potentially important distinctions; nevertheless, the analyses of these categories yielded some interesting results and provide direction for future research.

The major research design classification used in this meta-analysis was based on how offenders were assigned to the treatment and comparison groups. Any number of features can influence perceptions of study quality; however, Bangert-Drowns, Wells-Parker, and Chevillard (1997) found that ratings of study quality were most strongly linked to methods of subject assignment. Consequently, the studies were divided as follows: (a) studies in which preexisting group differences would not be expected (random assignment), (b) studies in which equivalence was not assured, but where there were no obvious reasons for group differences (incidental assignment), and (c) studies in which differences would be expected (dropouts, assignment based on need).

Random assignment is a strong method for equating groups, but no method guarantees equivalence. It is always possible for the groups to differ through purely random processes (see Marques, 1999). Random assignment studies have the important feature, however, that even if a particular comparison finds the treated and untreated groups to differ purely due to chance, such differences would "wash out" as more random assignment studies are completed.

In contrast, mismatched groups are expected when the decision to attend treatment is made by the offender, the program, or both (e.g., comparison groups composed of treatment refusers, dropouts, or offenders considered by staff to be high need or low need). Such designs have relatively little probative value because the treatment and comparison groups would be expected to differ on any number of relevant risk factors (e.g., offender compliance, impulsivity, or attitudes toward authority), and such differences would not "wash out" as the number of such studies increased.

Most sex offender treatment outcome studies have attempted to create equivalent groups through matching on risk-related variables rather than random assignment. For example, some studies created comparison groups from offenders sentenced in the same jurisdiction prior to the implementation of the treatment program, or from offenders explicitly matched on risk-related variables. When there was no obvious reason to expect pretreatment group differences, these studies were considered "incidental assignment."

Some studies led to considerable debate as to their appropriate classification. Such classification difficulties can partly be attributed to the subjectivity of the global coding system, but they also reflect legitimate differences concerning judgments of study quality. There are no perfect studies in the sex offender field, nor in any other field for that matter. What is a tolerable lapse to one expert could be a fatal flaw to another. One of the objectives of the Collaborative Outcome Data Project, and this current report, is to promote professional debate concerning the relative quality of treatment outcome studies for sex offenders.

Not only is it worth distinguishing between different types of research designs, it is also important to distinguish between different types of treatment. Many of the studies reviewed by Furby et al. (1989) examined treatments delivered in the 1960s and 1970s that would not meet current standards of practice. Defining standards for effective treatment is difficult given the active debate about whether treatment works at all. Nevertheless, sex offender treatment has evolved during the last 20 years, guided partly by theory (e.g., Laws, 1989) and partly by the research on "what works" for offenders in general (e.g., Andrews et al., 1990; Lösel, 1995). Programs considered most likely to be effective are those that target a range of criminogenic needs (e.g., attitudes tolerant of sexual assault, lifestyle impulsivity, deviant sexual interests) by teaching relevant skills in a manner appropriate for the learning style and receptivity of the offender.

Most treatment programs, however, were not sufficiently documented to allow ratings of treatment quality to be made with any confidence. Consequently, the current meta-analysis examined a simple distinction between older and current treatments. The criteria used for "current" were any treatment still being offered, or any cognitive–behavioral treatment delivered after 1980. Although a number of different approaches are potentially effective, sex offender treatment providers have increasingly put their faith in some version of cognitive–behavioral treatment.

The current study is the first report of an ongoing project that aims to include all credible studies of sex offender treatment. The studies considered in this report were those identified as of May 2000, in which sex offenders receiving psychological treatment were compared to sex offenders receiving no treatment or a form of treatment judged to be inadequate or inappropriate. Future reports will examine medical and drug treatments, and the effects of treatment on different types of sex offenders (e.g., child molesters, rapists).

METHOD

Selection of Studies

Computer searches of both PsycLIT and the National Criminal Justice Reference System (NCJRS) were conducted using the following key terms: sex(ual)

offender, rape, rapist, child molester, pedophile, pedophilia, exhibitionist, exhibitionism, sexual assault, incest, voyeur, frotteur, indecent exposure, sexual deviant, paraphilia(c), and treatment, outcome, recidivism, recidivist, recidivate, reoffend, reoffense, relapse, and failure. Additional articles were sought through the examination of the reference lists of the collected articles and those of review articles in this area. Finally, letters were sent to 30 established researchers in the field of sex offender recidivism requesting overlooked or as-yet unpublished manuscripts or data.

To be included in the present analysis, a study had to compare the recidivism rates (sexual or general) of a sample of treated sex offenders with a comparison group of sex offenders. The same recidivism criteria must have been used for both groups, and the recidivism rates must have been reported for approximately the same follow-up period. The combined sample must have been at least 10 (5 offenders in each group). The programs must have provided predominantly psychological treatment (e.g., group therapy, aversive conditioning), although a few studies included some offenders who received various forms of medication.

As of May 2000, our search yielded a total of 43 usable studies of psychological treatment for sex offenders (see Table I). This should not be considered an exhaustive list, but it contains more than twice the number of studies of the largest previous review (Gallagher et al., 1999, 20 studies of psychological treatment).

When different articles reported findings based on the same sample of offenders (or overlapping samples), the results were coded from the article reporting on the largest sample size and longest follow-up period. The same sample, however, could provide more than one design (e.g., random assignment as well as dropouts vs. completers). Table I provides descriptive information for the strongest design for each sample.

For one program (RTC Ontario), the results from two reports were combined (averaged) because the committee could not agree on which report to include (Davidson, 1984; Looman, Abracen, & Nicholaichuk, 2000). Two studies (Abracen, Looman, & Nicholaichuk, 1999; Mander, Atrops, Barnes, & Munafo, 1996) were not included following personal communications from the authors indicating unresolved anomalies in their data (Barnes, November 17, 1999; Looman et al., 2000). Different information was often drawn from different sources (e.g., recidivism information from a research study and the program description from an unpublished internal report). For 17 studies, the authors of the original reports coded the studies or provided supplemental information not available in the written documentation or both. Studies were classified as published ($k = 23$) or unpublished ($k = 20$) based on whether the recidivism data used for the meta-analysis were published in a book or journal.

Most of the studies were based on American ($k = 21$) or Canadian samples ($k = 16$), with five studies from the United Kingdom and one from New Zealand. The studies were mostly recent (median publication year = 1996), with 10 (23%) produced in 1999 or later (1977–2000). Total sample size ranged from 14 to 627

(median of 155). Most of the studies focused on adult male sex offenders. Four studies specifically examined adolescent sex offenders. Only one study (Barnes & Peterson, 1997) indicated that their sample contained female offenders (<5% of their total sample) and none of these women had sexually recidivated.

Of the 43 programs, 23 were offered in institutions, 17 in the community, and 3 in both settings. The major sponsor of the programs was corrections ($k = 26$), followed by health ($k = 9$) and private clinics ($k = 8$). Treatment was delivered between 1965 and 1999, with approximately 80% of the offenders receiving treatment after 1980. Most studies examined specialized sex offender treatment programs, although three studies examined the response of sex offenders to programs designed for general offenders (Borduin, Henggeler, Blaske, & Stein, 1990; Nutbrown & Stasiak, 1987; Robinson, 1995).

Recidivism was defined by reconviction in 8 studies and rearrest in 11 studies. Twenty studies used broad definitions of recidivism that included parole violations, readmissions to institutions, or unofficial community reports or all of these. The most common source of recidivism information was national criminal justice records ($k = 26$) followed by state/provincial records ($k = 19$). Additional sources of information (e.g., child protection records, self-reports) were used in 9 studies. For 6 studies, the source of the recidivism information was not specified. Thirteen studies reported only sexual recidivism, 5 studies reported only general (any) recidivism, and 25 studies reported on both sexual and general recidivism.

The average follow-up periods ranged from 12 months to 16 years, with a median of 46 months for both the treatment and comparison groups.

Coding Categories

Studies were assigned to one of the following six research-design categories: (a) random assignment (versus no treatment or alternate treatment); (b) incidental assignment (versus no treatment or alternative treatment); (c) any treatment attendance (including dropouts) versus treatment refusers; (d) treatment completers versus dropouts; (e) dropouts versus refusers; and (f) assignment based on need (e.g., treatment given to those assessed as requiring treatment).

In most cases, the "refusers" were offenders in the comparison group who were assessed and offered treatment, but declined to participate. Comparison groups could also be coded as "refusers" when the treated sample comprised volunteers who were heavily screened before being admitted.

The 17 incidental assignment studies included comparison groups drawn from the following sources: offenders released before the implementation of the treatment program ($k = 5$), offenders matched from archives of criminal history records ($k = 3$), offenders receiving an earlier version of the treatment program

Table I. Studies Used in the Meta-Analysis

ID	Study	Type of treatment	Strongest design	Sample size Treatment	Sample size Comparison	Years of follow-up	Sexual recidivism rates Treatment	Sexual recidivism rates Comparison	Odds ratio (sexual)	Any recidivism Treatment	Any recidivism Comparison	Odds ratio (any)
1	Alberta Hospital	Cog.–beh.[c]	Dropouts	194	157	5	.04	.08	0.43	.16	.34	0.39
2	Allam (1998, 1999)	Cog.–beh.[c]	Incidental—treatment not condition of sentence & matched on risk	153	74	1	.03	.08	0.39	.13	.42	0.21
3	Bakker, Hudson, Wales, & Riley (1999)	Cog.–beh.[c]	Incidental—prior to program implementation	238	283	8	.09	.19	0.45			
4	Barnes & Peterson (1997)	Cog.–beh.[c]	Volunteers/Refusers	147	138	3	.03	.09	0.39	.12	.26	0.40
5	Berlin et al. (1991)	Psychoth.	Dropouts	257	206	5	.05	.15	0.28			
6	Bluglass (1980)	Psychoth.	Need	43	57	5				.42	.33	1.43
7	Borduin et al. (1990, 2000)[a]	Systemic[c]	Random	24	24	8	.13	.42	0.22	.25	.88	0.08[b]
8	Bremer (1992)	Psychoth.	Dropouts	66	27	—	.18	.11	1.61			
9	Clearwater	Cog.–beh.[c]	Incidental—matching from archives	245	218	5	.13	.24	0.47			
10	CS/RESORS (1991)	Cog.–beh.[c]	Incidental—variable program integrity	335	181	3	.05	.04	1.19			
11	Dwyer	Cog.–beh.[c]	Dropouts	125	55	8	.06	.16	0.35	.09	.20	0.39
12	Florida	Psychoth.	Need	39	199	1	.05	.05	1.20	.13	.18	0.74
13	Guarino-Ghezzi & Kimball (1998)[a]	Cog.–beh.[c]	Incidental—Administrative alternate treatments	33	25	1	.00	.04	0.24	.30	.48	0.48
14	Hall (1995a)	Cog.–beh.[c]	Volunteers/Refusers	24	6	1	.00	.17	0.07	.21	.50	0.28
15	Hanson, Steffy, & Gauthier (1992, 1993)	Psychoth.	Incidental—prior to program implementation; administrative	102	89	16	.37	.33	1.22	.63	.61	1.09
16	Hersh, Baldwin, & Gray-Little (1999)	Cog.–beh.[c]	Dropouts	285	100	—				.38	.59	0.43
17	Huot (1999)	Cog.–beh.[c]	Volunteers/Refusers	92	159	7	.16	.19	0.82	.14	.47	0.19
18	Kramer (1985)	Behavioral	Incidental—change in program	37	19	—	.05	.32	0.15			
19	Lab et al. (1993)[a]	Cog.–beh.[c]	Need	46	109	2	.02	.04	0.77	.26	.20	1.41
20	Lindsay	Cog.–beh.[c]	Incidental—longer versus shorter probation orders	7	7	—	.00	.57	0.05			

#	Study	Treatment	Comparison group		Treatment *n*	Comparison *n*						
21	La Macaza	Cog.–beh.[c]	Incidental—prior to program implementation	3	65	56	.06	.21	0.41	.14	.47	0.19
22	Marques	Cog.–beh.[c]	Random	5	190	225	.16	.16	1.06	.65	.59	1.30
23	Marshall & Barbaree (1988)	Cog.–beh.[c]	Dropouts	4	68	58	.13	.34	0.30			
24	Marshall et al. (1991)	Cog.–beh.[c]	Incidental—change in program	7	17	23	.24	.35	0.50	.10	.34	0.22
25	McGrath et al. (1998)	Cog.–beh.[c]	Incidental—alternate treatment	5	71	32	.01	.16	0.11			
26	McGuire (2000)	Cog.–beh.[c]	Dropouts	—	54	14	.00	.07	0.08	.04	.07	0.43
27	Missouri	Cog.–beh.[c]	Dropouts	4	105	156	.05	.13	0.36	.23	.41	0.43
28	Nutbrown & Stasiak (1987)	Psychoth.	Incidental—administrative	3	59	11				.20	.55	0.22
29	Perkins (1987)	Cog.–beh.	Incidental—various methods	—	62	12	.32	.17	3.38	.48	.57	0.72
30	JJ Peters	Psychoth.	Random	—	148	83	.14	.07	1.96	.55	.60	0.82
31	Pinel	Cog.–beh.[c]	Dropouts	6	117	55	.25	.24	1.05	.37	.56	0.45
32	Procter (1996)	Cog.–beh.[c]	Incidental—prior to program implementation	4	54	54	.07	.15	0.49			
33	Rattenbury (1986)	Unknown	Need	6	69	69	.22	.14	1.61	.39	.36	1.13
34	Rice et al. (1991)	Behavioral	Need	6	51	85	.51	.28	2.61	.65	.54	1.54
35	Robinson (1995)	Cog.–beh.[c]	Random	6	189	46				.30	.46	0.51
36	RHC Pacific	Cog.–beh.[c]	Need	2	38	29	.08	.00	5.82	.34	.21	1.19
37	RTC Ontario 1976/1989	Cog.–beh.[c]	Incidental—matched from archives/prior to program	5	95	95	.26	.32	0.78			
38	Saskatchewan	Cog.–beh.[c]	Incidental—matched from archives	2	26	35	.12	.03	3.43			
39	Twin Rivers	Cog.–beh.[c]	Volunteers/Refusers	3	209	97	.02	.08	0.23	.15	.25	0.53
40	Walker (2000)	Systemic[c]	Dropouts	2	44	62	.00	.06		.00	.11	0.08
41	Warkworth	Cog.–beh.[c]	Volunteers/Refusers	3	312	85	.06	.14	0.92	.13	.15	0.78
42	Washington	Mixed	Volunteers/Refusers	7	321	306	.11	.13	0.75	.20	.52	0.23
43	Worling & Curwen (1998)[a]	Systemic[c]	Incidental—referred for assessment only	—	85	46	.12		0.87			

Note. Cog.–beh. – Cognitive–behavioral treatment.

[a] Denotes a study of adolescent sex offenders.

[b] Any recidivism from Borduin et al. (1990), *n* = 16.

[c] Denotes a current treatment.

($k = 2$), and offenders who received no treatment ($k = 2$) or treatment judged to be lower in quality ($k = 3$) due to administrative reasons, such as program unavailability or insufficient time remaining in their sentence.

The distinctions between designs were usually clear (see Discussion of reliability below). The following are some of the more controversial decisions and are described to help illustrate the boundaries of the coding categories. Robinson (1995) randomly assigned offenders to treatment or a waiting list condition. The assignment was not pure, however, because 27% (199 of 740) of the waiting list group subsequently received treatment and were placed in the treatment sample ($n = 3,531$). The majority of the committee considered this to be a random assignment study because all offenders initially expressed interest in participating and the primary reasons for switching from the comparison group to the treatment group was the lengthy sentence being served by the offender (which, if anything, should decrease treatment effects).

Marshall and Barbaree (1988) compared a treatment group to offenders who expressed interest in receiving treatment, were assessed at the community clinic, but did not attend treatment because they "lived too far away." This study was coded as "refusers/drop-outs" because the offenders' motivation may have been a factor influencing how far away is "too far."

In the study by McGrath, Hoke, and Vojtisek (1998), the offenders could choose between specialized group treatment and individual treatment. Although the offenders had a choice of which treatment they received, the study was coded as incidental assignment because the offenders' choices appeared to be based mainly on administrative reasons (e.g., they were already involved with a therapist that they liked).

In Allam's study, the comparison group included offenders who were assessed in presentence reports as appropriate for treatment, but treatment was not made a condition of probation so they did not attend (Allam, 1999). The study was coded incidental assignment because it appears that the judges' determination of who received treatment was not based on the offenders' risk level or their willingness to attend treatment. The untreated offenders were slightly at higher risk than the treated offenders were based on an actuarial measure of risk for sexual offence recidivism (Structured Anchored Clinical Judgment, see Hanson & Thornton, 2000).

The comparison subjects in Worling and Curwen (1998) were adolescent offenders referred by other treatment agencies for specialized assessment, not treatment. This study was considered incidental assignment because the program did not determine who would receive treatment, although need for treatment may have been determined by the referring agencies. The effect size calculated in the current study, however, was substantially smaller than that reported by Worling and Curwen because the present analysis included treatment dropouts with the treatment group, not the comparison group.

Some of the committee recommended completely removing the study by Nicholaichuk et al. (2000) because of a known artefact in the data (Hanson & Nicholaichuk, 2000). The majority decision was to retain the study as incidental assignment, but to use only the data in which the chances of corruption were small.

The type of treatment was initially categorized as (a) cognitive–behavioral ($k = 29$), (b) behavioral ($k = 2$), (c) systemic ($k = 3$), or (d) other psychotherapeutic ($k = 7$). In two studies, the nature of the treatment was unknown. The studies were then grouped as "current" or "non-current." Current treatments met one of the following criteria: (a) any treatment currently being offered (1998–2000), either at the original site, or at other sites via treatment manuals; or (b) all cognitive–behavioral treatments given to offenders after 1980. This definition of "current" identified 27 studies of cognitive–behavioral treatment and three studies of systemic treatment (2 of adolescents; 1 of incest offenders). Systemic treatment aims to change offending behavior by addressing the needs of the family and other social systems influencing the individual, such as school, peers, and courts (Swenson, Henggeler, & Schoenwald, 2001).

Coding Procedure

All published and unpublished documentation was collected for each study. One of the three trained raters (psychology graduate students) then coded the study using a standardized coding manual (containing over 250 variables, such as sample size, follow-up period, methods of treatment used). The authors of the original study were contacted for clarification when insufficient information was available concerning crucial variables (e.g., method of subject assignment). Five studies were coded by the committee directly from the original data (Clearwater, CS/RESORS, Hanson, Alberta Hospital, Pinel). As well, the authors of four of the original reports coded all the data except for the design classification, which was done by the committee. The initial codings were reviewed by one of the committee members (R. K. H.). Most of the disagreements involved inattention or clerical errors, which were immediately resolved when the relevant information was identified. Even when the initial two raters agreed, potentially controversial decisions were referred to the committee. In the few cases where consensus could not be reached, the final coding was based on majority opinion of the committee.

To evaluate the reliability of the design classification, an independent rater identified the strongest design for each study (random, incidental, refusers, dropouts, or need, in that order). The rater, a graduate student in psychology, had no previous involvement with this project but had prior experience with sex offender research. The raters' judgments were identical for 35 of the 45 ratings

(78% agreement, $\kappa = .71$). This level of rater reliability is similar to that found in other meta-analyses (Bangert-Drowns et al., 1997).

Whenever possible, data were coded to fit the research design that minimized the likelihood of preexisting differences between the treatment and comparison groups. Consequently, the recidivism rates reported in Table I do not always correspond to those reported in the original articles. If sufficient information was available, dropouts were included in the treatment group. If survival data were presented, recidivism rates were taken directly from the graph for a standard follow-up period. This time period was somewhat before the end of the longest possible follow-up period because survival curves are unstable when a small number of offenders ($n < 20$) remain in the sample.

Index of Treatment Effectiveness

The basic outcome data were 2×2 tables containing the recidivism outcomes of the treatment and comparison groups. Several different statistics, such as rate ratios and phi, can be used to summarize 2×2 tables for the purpose of meta-analysis. Following the recommendation of Fleiss (1994), the index selected was the odds ratio, which is defined as follows: $(\text{recid}_t/\text{nonrecid}_t)/(\text{recid}_c/\text{nonrecid}_c)$, where recid_t is the number of recidivists in the treatment group, nonrecid_t is the number of nonrecidivists in the treatment group, recid_c is the number of recidivists in the comparison group, and nonrecid_c is the number of nonrecidivists in the comparison group.

In contrast to other common effect size indicators, the odds ratio is relatively unaffected by arbitrary design features, such as the proportion of offenders in the treatment and comparison groups, or the overall recidivism rate (Fleiss, 1994). As recommended by Fleiss, a value of 0.5 was added to each cell to minimize bias and to permit the analysis of tables containing empty cells.

Values of the odds ratio can range from very small (e.g., $<.01$) to very large (e.g., >100) with values of 1.0 indicating no difference between the groups. Small values of the odds ratio indicate treatment effectiveness, that is, lower recidivism rates in the treatment than comparison groups. When the recidivism base rate is low, the odds ratio approximates the rate ratio. For example, an odds ratio of 0.70 can be interpreted as follows: for every 100 untreated sex offenders who recidivate, only 70 treated sex offenders will recidivate.

Summarizing Findings

Standard meta-analytic procedures were used to summarize the results across studies (Hedges, 1994; Hedges & Olkin, 1985). Variability across studies was indexed by the Q statistic: $Q = \sum w_i (\text{LOR}_i - \text{LOR}_m)^2$, where LOR_i is the natural

log of the odds ratio for each sample, w_i is the weight for each sample (equal to the inverse of its variance − SE^2), and LOR_m is the weighted grand mean ($\sum w_i LOR_i / \sum w_i$). The Q statistic is distributed as χ^2 with degrees of freedom equal to $k - 1$, where k is the number of studies.

Statistical tests were conducted on the log of the odds ratio [$\log(n_{11}n_{22}/n_{12}n_{21})$] because the standard error of the log odds ratio is easily defined as $(1/n_{11} + 1/n_{12} + 1/n_{21} + 1/n_{22})^{1/2}$ (n_{11}, n_{12}, n_{21}, and n_{22} are the cells of a 2×2 table). The reported statistics, however, were transformed back into odds ratios for ease of interpretation.

RESULTS

The 43 studies examined a total of 5,078 treated sex offenders and 4,376 untreated sex offenders. Averaged across all studies, the sexual offence recidivism rate was lower for the treatment groups (12.3%) than the comparison groups (16.8%, 38 studies, unweighted average). A similar pattern was found for general recidivism, although the overall rates were predictably higher (treatment 27.9%, comparison 39.2%, 30 studies). These recidivism rates were based on an average 46-month follow-up period using the variety of recidivism criteria reported in the original studies.

Sexual Recidivism

Table II presents the meta-analytic summary of the results for sexual recidivism. Averaged across all types of treatment and all research designs, there was a small advantage for the treated versus the untreated offenders (OR = 0.81). Given that the confidence interval (0.71–0.94) did not include "1.0," the effect was statistically significant. There was, however, considerable variability across studies, as indicated by a large Q statistic (145.02, $df = 37$, $p < .001$). The treatment effect was much stronger in the unpublished studies (OR = 0.65, 95% CI of 0.52–0.81, $k = 17$) than in the published studies (OR of 0.95, 95% CI of 0.79–1.15, $k = 21$).

Three random assignment studies used sexual recidivism as the outcome variable. These diverse studies produced diverse results. Borduin, Schaeffer, and Heilblum (2000) found a significant effect of multisystemic treatment for adolescent sex offenders (OR = 0.22). Romero and Williams (1983) found that weekly, unstructured group psychotherapy was associated with a nonsignificant increase in sexual recidivism (OR = 1.90). Marques (1999) found no difference for offenders receiving cognitive–behavioral treatment and the comparison group (OR = 1.09). Overall, the odds ratio associated with the three random assignment studies was 1.03 (95% CI of 0.67–1.59), with more variability between the studies than would be expected by chance ($Q = 6.36$, $df = 2$, $p < .05$).

Table II. Treatment Effectiveness for Reducing Sexual Offence Recidivism

Studies	Odds ratio	95% CI	Q	n (k)
All	0.81	0.70/0.93	145.02***	8,164 (38)
Research designs				
Random assignment	1.03	0.67/1.59	6.36*	694 (3)
Incidental assignment	0.62	0.50/0.77	28.98*	2,948 (17)
Completers vs. dropouts	0.47	0.36/0.61	21.50	2,732 (18)
Any attendance vs. refusers	0.90	0.71/1.14	24.93**	2,892 (11)
Dropouts vs. refusers	1.67	0.91/3.06	2.69	353 (4)
Assignment based on need	3.10	2.15/4.48	11.10*	1,130 (6)
Treatment comparisons				
Current treatments				
Adults	0.61	0.48/0.76	21.17*	2,779 (12)
Adolescents	0.50	0.22/1.13	2.53	237 (3)
Total	0.60	0.48/0.75	23.89*	3,016 (15)
Institutional (current)				
Adults	0.62	0.48/0.80	12.31*	1,771 (6)
Adolescents	0.24	0.01/6.24		58 (1)
Total	0.61	0.48/0.79	12.62*	1,829 (7)
Community (current)				
Adults	0.57	0.34/0.95	8.78	1,008 (6)
Adolescents	0.53	0.23/1.22	2.32	179 (2)
Total	0.56	0.36/0.86	11.12	1,187 (8)
Non-current treatment	1.19	0.77/1.86	8.19	626 (5)

Note. 95% CI is the 95% confidence interval for the odds ratio, Q is a measure of between study variability (small values indicate less variability), n is the total number of offenders, and k is the number of studies.
$^{*}p < .05.$ $^{**}p < .01.$ $^{***}p < .001.$

On average, the 17 studies using incidental assignment were associated with significant reductions in sexual recidivism (OR = 0.62, 95% CI of 0.50–0.77) with more variability than would be expected by chance ($Q = 29.0$, $df = 16$, $p < .05$).

Those offenders who dropped out of treatment had consistently higher sexual recidivism rates than those who completed treatment (OR of 0.47). The effect was consistent across the 18 studies ($Q = 21.50$, $df = 17$, $p > .10$).

Offenders who refused treatment, however, did not appear to be at higher risk than offenders who attended any treatment (including dropouts). The odds ratio was 0.90 with significant variability across studies ($Q = 24.93$, $df = 10$, $p < .01$). The four studies that directly compared dropouts to refusers found higher sexual recidivism rates among the dropouts than the refusers (OR of 1.67), but the effect was not statistically significant.

Offenders referred to treatment based on perceived need had significantly higher sexual recidivism rates than the offenders considered not to need treatment (OR of 3.10). The only study not finding this pattern was Lab, Shields, and Schondel (1993), in which treatment was provided to offenders with the lowest perceived risk. When this outlier was removed the effect for need-based assignment increased to 3.38 with no significant variability ($Q = 8.92$, $df = 4$, $p > .05$).

The remarkably consistent effects found when the comparison groups included dropouts or those assessed as not needing treatment suggests that the results of these studies are more related to the method of subject selection than to the treatment received. Consequently, comparisons between treatments were conducted using only the incidental and random assignment studies. These two types of studies were combined because there were only two random assignment studies that examined current treatment: Borduin et al. (1990, 2000), who found a significant treatment effect, and Marques (1999), who did not. It should be noted that the Marques (1999) data used in the current study were the preliminary 1995 results; the final report from this study has yet to be released.

On average, the 20 incidental/random studies were associated with an overall treatment effect (OR = 0.69; 95% CI of 0.56–0.84), with substantial variability across studies ($Q = 39.63, df = 19, p < .01, n = 3,642$). For the 15 studies of current treatment, the odds ratio was 0.60 (95% CI of 0.48–0.76), with significant variability ($Q = 23.89, df = 14, p < .05$). The weighted average across these 15 studies corresponds to a sexual recidivism rate of 9.9% for the treatment groups ($n = 1,638$) compared to a rate of 17.4% for the comparison groups ($n = 1,378$, medians of 9.2 and 15.6%).

The treatment effect remained significant when considering only studies in which dropouts were explicitly included with the treatment group (OR of 0.56, 95% CI of 0.35–0.91, $Q = 12.34, df = 6, p > .05, k = 7$). In contrast, noncurrent treatments had no effect on reducing sexual recidivism (OR of 1.19; $k = 5$).

Studies that compared a current treatment to an alternate form of treatment tended to find larger treatment effects (OR of 0.28, 95% CI of 0.15–0.54, 6 studies, $n = 708; Q = 3.62, ns$) than did studies that compared current treatments to an untreated comparison group (OR of 0.64, 95% CI of 0.51–0.81, 10 studies, $n = 2,753; Q = 15.71, ns$).

Current treatments appeared to be equally effective for adults (OR of 0.61) and adolescents (OR of 0.50). None of the effects were statistically significant for the adolescent programs, however, due to the small sample size (total $n = 237; k = 3$).

Institutional treatment (OR of 0.52) and community treatment (OR of 0.56) were both associated with reductions in sexual recidivism. There was relatively little variability across the studies, with the amount of variability in the community treatment studies being no more than would be expected by chance ($Q = 11.12, df = 7, p > .05$).

The effects were similar for institutional programs treating adults (OR of 0.62) and adolescents (OR of 0.24), although it was difficult to make a comparison because only one study examined an institutional program for adolescents (Guarino-Ghezzi & Kimball, 1998). Current community treatments were associated with reductions in sexual recidivism for adults and adolescents (OR of 0.57 and 0.53, respectively). Again, the effect for the adolescents was not significant due to low sample size ($n = 179, k = 2$).

Table III. Treatment Effectiveness for Reducing General (Any) Recidivism

Studies	Odds ratio	95% CI	Q	n (k)
All	0.56	0.50/0.64	120.08***	6,075 (31)
Research designs				
Random assignment	0.92	0.69/1.22	10.61*	897 (4)
Incidental assignment	0.52	0.40/0.68	34.70***	1,176 (10)
Completers vs. dropouts	0.38	0.30/0.47	8.56	2,202 (14)
Any attendance vs. refusers	0.47	0.39/0.58	48.26***	2,274 (8)
Dropouts vs. refusers	0.85	0.42/1.73	0.18	123 (2)
Assignment based on need	1.30	0.94/1.82	2.26	834 (6)
Treatment comparisons				
Current treatments				
Adults	0.59	0.45/0.78	33.00***	1,101 (5)
Adolescents	0.35	0.13/0.91	2.05	74 (2)
Total	0.57	0.44/0.74	36.12***	1,175 (7)
Institutional (current)				
Adults	0.82	0.60/1.13	15.76***	771 (3)
Adolescents	0.48	0.17/1.39		58 (1)
Total	0.79	0.58/1.07	16.66***	829 (4)
Community (current)				
Adults	0.21	0.12/0.37	0.01	330 (2)
Adolescents	0.08	0.01/0.75		16 (1)
Total	0.20	0.12/0.35	0.73	346 (3)
Non-current treatment	0.84	0.63/1.12	13.11*	898 (7)

Note. 95% CI is the 95% confidence interval for the odds ratio, Q is a measure of between study variability (small values indicate less variability), n is the total number of offenders, and k is the number of studies.
$*p < .05$. $***p < .001$.

General Recidivism

The treatment effects for general recidivism were similar to the effects for sexual recidivism (See Table III). Across all studies ($n = 31$), treated offenders had significantly lower general recidivism rates than untreated offenders (OR of 0.56), with substantial variability across studies ($Q = 120.08$, $df = 30$, $p < .001$). The 16 published studies (OR of 0.59, 95% CI of 0.49–0.70) and the 15 unpublished studies (OR of 0.54, 95% CI of 0.46–0.64) showed similar results.

Two of the four random assignment studies found significant treatment effects. Robinson (1995) found a cognitive skills training program to be associated with a significant reduction in general offending (OR of 0.51). Similarly, Borduin et al. (1990) found that multisystemic treatment reduced general recidivism among adolescents (OR = 0.08, $p < .05$). The Borduin et al. results need to be interpreted cautiously, however, because they were based on a sample of only 16 subjects (8 per group); the general recidivism rate was not reported in the larger extension of the study (Borduin et al., 2000). Romero and Williams (1983), however, found no significant effect for unstructured group psychotherapy for adult offenders (OR of 0.82) nor did Marques (1999) find any effect for a cognitive–behavioral program

(ORs of 1.30). Overall, the random assignment studies were associated with a non-significant odds ratio of 0.92 (95% CI of 0.69–1.22), with significant variability ($Q = 10.61, df = 3, p < .05$).

The 10 incidental assignment studies were associated with a reduction in general recidivism (OR of 0.52, 95% CI of 0.40–0.68) with significant variability ($Q = 34.70, df = 9, p < .001$).

Offenders who completed treatment had consistently lower general recidivism rates than those who failed to complete treatment (OR of 0.38) and the effect was consistent across studies ($Q = 8.56, df = 13, p > .75$). In contrast to the findings for sexual recidivism, however, those who refused sex offender treatment were more likely to recidivate with any offence than those who attended any treatment (OR of 0.47, with significant variability). The two studies that compared dropouts with refusers on general recidivism found similar rates for both groups (both relatively high risk). Studies in which offenders were assigned to sex offender treatment based on perceived need showed similar rates of general recidivism for both the treated and untreated groups (OR of 1.30, 95% CI of 0.94–1.82, $Q = 2.26, df = 5, p > .50$).

When the analysis was restricted to the incidental and random-assignment studies, the odds ratio was 0.69 (95% CI of 0.57–0.84, $k = 13$). The effect was higher for current treatments (OR of 0.57; 95% CI of 0.44–0.74; $k = 7$), than for noncurrent treatment (OR of 0.84, 95% CI of 0.63–1.12, $k = 7$). A weighted average for the 7 studies of current treatments resulted in a general recidivism rate of 32.3% for the treated offenders ($n = 709$) and 51.3% for the comparison groups ($n = 466$) (medians of 25 and 46%, respectively).

As with sexual recidivism, studies that compared a current treatment to an alternate form of treatment tended to find larger treatment effects (OR of 0.28, 95% CI of 0.14–0.57, 3 studies, $n = 177$) than studies that compared current treatments to an untreated comparison group (OR of 0.64, 95% CI of 0.48–0.85, 4 studies, $n = 998$).

Treatment reduced general recidivism significantly for both adults (OR of 0.59) and adolescents (OR of 0.35).

Current community treatments appeared to have a stronger effect on general recidivism (OR of 0.20; 95% CI of 0.12–0.35; $k = 3$) than did treatment provided in institutions (OR of 0.79; 95% CI of 0.58–1.07; $k = 3$). This difference should be interpreted cautiously, however, given the small number of studies and the significant variability in the effectiveness of the institutional treatments ($Q = 16.66, df = 3, p < .001$). The relative effectiveness of community versus institutional treatment appeared similar for adults and adolescents, but there was only one study of institutional treatment and one study of community treatment for adolescents.

An important question is the extent to which sex offender specific treatment is effective in reducing general recidivism. The four studies (1 random; 3 incidental) of current, sex offender specific treatment for adults found a significant reduction

in general recidivism (OR of 0.61, 95% CI of 0.45–0.82, $Q = 32.79$, $df = 3$, $p <$ .001, $n = 866$).

DISCUSSION

The current review, like previous quantitative reviews (Gallagher et al., 1999; Hall, 1995b), found that the recidivism rates of treated sex offenders were lower than the recidivism rates of untreated sex offenders. Given the large numbers in the current study (more than 9,000 offenders in 43 studies), this pattern of results cannot be seriously disputed. What can be disputed are the reasons for the group differences. Did the treatment reduce the offenders' recidivism rates, or were the observed differences produced by unintended consequences of the research designs? We believe that the balance of available evidence suggests that current treatments reduce recidivism, but that firm conclusions await more and better research.

The strongest research designs are those in which offenders are randomly assigned to treatment, but there are very few of such studies. Borduin et al. (1990, 2000) found multisystemic treatment to be effective with adolescent sex offenders, but this form of treatment has yet to be evaluated in other settings, and is difficult to apply to adult sex offenders. Robinson (1995) found that a cognitive skills training program reduced general recidivism among sex offenders. Robinson's (Robinson, 1995) findings, however, are best considered part of the already well-established research on "what works" with general offenders (e.g., Lösel, 1995). There is only one random assignment study examining a current sex offender specific treatment (Sex Offender Evaluation and Treatment Project SOTEP; Marques, 1999), which, so far, has not found a positive effect of treatment. Rather than limit the entire sex offender treatment debate to the strengths and weaknesses of the SOTEP study, the Collaborative Project considered research studies using methods other than random assignment.

Studies comparing treatment completers to dropouts consistently found higher recidivism rates for the dropouts, regardless of the type of treatment provided. Even in studies where there was no difference between the treatment group and the untreated comparison groups, the treatment dropouts did worse. Consequently, treatment dropout studies cannot be used to evaluate the effectiveness of treatment (except in the discouraging case where dropouts do better than completers).

Dropouts can be high risk to reoffend for a number of reasons. They are likely to have preexisting characteristics associated with recidivism risk (e.g., youth, impulsivity, and unstable lifestyles; Wierzbicki & Pekarik, 1993) and the factors motivating treatment termination are often themselves correlated with recidivism (e.g., offender hostility toward authority, noncompliance). It is also possible that

interrupted treatment makes offenders worse. The initial stages of treatment can introduce offenders to deviant role models, cognitive distortions, and a wide range of novel, sexually deviant fantasies and behaviors. Although the mechanisms have yet to be identified, the consistent findings suggest that failure to complete sex offender treatment is a reliable and robust predictor of recidivism.

A somewhat surprising finding was that offenders who refused treatment were not at higher risk for sexual recidivism than offenders who started treatment. Treatment refusers, however, were at relatively high risk for general recidivism. Several explanations for this pattern of results are available. Some offenders may realistically conclude that they do not require treatment. Refusing treatment may not be a marker of sexual deviance; instead, it may be associated with a generally noncooperative, antisocial lifestyle. There was, however, significant variability across studies and further research is required to determine when, if ever, treatment refusal is associated with an increased risk. The current results, however, are a challenge to evaluators who routinely use "treatment refusal" as a poor prognostic indicator.

Studies that compared sex offenders who "needed" treatment to less needy offenders consistently found worse outcomes for the treatment group. It appears that evaluators are better able to identify high risk offenders than to change them. The results of need-based assignment studies are difficult to interpret because their outcome is determined by the competing influences of the need assessment and treatment.

Most of the evidence for treatment effectiveness came from incidental assignment studies. In these studies, there was no obvious, a priori reason that the offenders would be higher or lower in risk than the comparison group. Neither the offender nor the program determined who would receive treatment, and there were no other obvious factors that should bias the composition of the groups. The incidental assignment studies found an overall effect for treatment as well as variation across treatments. Current treatments (any treatment currently offered and cognitive–behavioral treatments offered since 1980) were associated with a significant reduction in both sexual and general recidivism whereas the older treatments were not.

When the random and incidental assignment studies were combined, current treatments were associated with reductions in both sexual recidivism (from 17.3 to 9.9%), and general (any) recidivism (from 51 to 32%). These reductions were not large, but they were statistically reliable and large enough to be of practical significance.

Although care was taken to identify and code studies in a manner that minimized preexisting group differences, it is always possible that the studies included hidden biases. The task for those concerned with the empirical evaluation of sex offender treatment programs is to identify and describe potential threats to validity. The importance of potential threats can then be debated and, wherever possible, empirically evaluated.

Consider, for example, a cohort design in which the comparison group included a sample of offenders from the 10 years prior to the implementation of the treatment program in 1990. One potential threat to validity is a cohort effect on recidivism rates: the chances of being reconvicted for sexual offence may be higher (or lower) in the 1990s than in the 1980s. Such cohort effects may be dismissed as trivial, but significant cohort effects have been documented. In the United Kingdom, the chances of a reported sexual offence resulting in a conviction dropped from 37% in 1981 to 17% in 1997 (Friendship & Thornton, 2001). Providing evidence that the outcome criteria did not undergo substantial change during the study period could increase the validity of cohort designs.

In many cases, however, the information required to examine threats to validity is not available. This lack of information can partly be attributed to a lack of diligence, but there is a more fundamental problem: uncertainty concerning what information needs to be collected. Even experts disagree as to the most important threats to validity in sex offender treatment outcome studies. Some guidance can be drawn from standard texts on research methodology (e.g., Campbell & Stanley, 1966), but the most salient factors vary with the research questions and the current state of knowledge in the field (Bangert-Drowns et al., 1997). A major aim of the Collaborative Outcome Data project is to promote debate and research concerning the essential requirements of sex offender treatment outcome studies.

It is interesting to note, however, that previous meta-analytic reviews have not found more methodologically rigorous research studies produce results substantially different from less rigorous studies (e.g., Redondo, Sánchez-Meca, & Garrido, 1999). Redondo et al.'s meta-analysis of European outcome studies for general offenders found that the highest quality designs were associated with the largest effect sizes (Redondo et al., 1999). In a meta-analysis of 191 remedial programs for intoxicated drivers, Bangert-Drowns et al. (1997) found that study quality was more related to the variability in the findings than to the effect size. As study quality increased, the results converged around the mean value.

The encouraging results of the current review raises questions as to why previous reviews have been inconclusive (e.g., United States General Accounting Office, 1996). First, the present results suggest that early reviewers had good reason to be sceptical: older treatment programs were associated with a slight (nonsignificant) increase in sexual recidivism (OR of 1.19). The treatment of sex offenders, however, has changed considerably since the 1970s. Studies of newer forms of treatment have only recently become available. Most of the studies in the current review were produced after 1995 and 23% were only available after 1999. All but one of the 1999/2000 studies were still unpublished at the time of this writing, resulting in larger treatment effects for the unpublished studies than the published studies.

Another reason that treatment effect may appear ephemeral is the low statistical power of the typical outcome study (Barbaree, 1997). Although the relative

reduction in recidivism rate was substantial (approximately 40%), the absolute reduction in recidivism rates was modest (7%), even among the better-designed studies of current treatments. Given this modest absolute reduction and a median sample size of 85 treated subjects, the average study would be expected to obtain significant results only 25% of the time (difference between .17 and .10; $\alpha = .05$, two-tailed; Cohen, 1988). Only 6 of the 15 current treatment studies were statistically significant in the original reports. Although some of the variability could be attributed to different samples and treatments, this number of nonsignificant findings is expected given low statistical power.

DIRECTIONS FOR FUTURE RESEARCH

This is the first report of the ongoing Collaborative Outcome Data project. There is much to learn about the effectiveness of treatment for sex offenders, and knowledge can only accumulate as new studies are made available. Clinicians can make a significant contribution to this project by organizing their records in a manner conducive to systematic evaluation (Hanson, 2000). Researchers can contribute by producing high-quality studies and making them available for systematic integration. Sample size matters less than the care taken to promote equivalence between the treatment and comparison groups. Ideally, this would involve random assignment. The accumulated evidence of small, nonsignificant studies (e.g., $n < 20$) can potentially yield valuable conclusions.

Although the meta-analysis provided evidence about the overall effectiveness of treatment, it provided little direction on how to improve current practise. The treatments that appeared effective were recent programs providing some form of cognitive–behavioral treatment, and, for adolescent sex offenders, systemic treatment aimed at a range of current life problems (e.g., family, school, peers). Further research is needed in order to make reliable distinctions between types of treatment and types of offenders. Different sex offenders would be expected to have different treatment needs; not all treatment would be expected to benefit all offenders (Marques, 1999).

ACKNOWLEDGMENTS

Looman, Timothy McGuire, Michael Miner, Terry Nicholaichuk, Richard Packard, Katherine Peterson, Jean Proulx, John Reddon, Marnie Rice, Greg Schiller, Marylee Stephenson, Lea Studer, and Thomas Wilson. This project would not have been possible without their help.

REFERENCES

Note. References marked with an asterisk indicate studies included in the meta-analysis. The name of the setting is given at the end of the reference when the studies that correspond to a single setting cannot be identified by the name of the first author.

Abracen, J., Looman, J., & Nicholaichuk, T. (1999). *Recidivism among treated sexual offenders and matched comparison subjects: Data from the Regional Treatment Centre (Ontario) post-1989 sample*. Unpublished manuscript, Regional Treatment Centre, Correctional Service of Canada, Kingston, Ontario.

Alexander, M. (1999). Sexual offender treatment efficacy revisited. *Sexual Abuse: A Journal of Research and Treatment, 11*, 101–116.

*Allam, J. (1998). *Community-based treatment for sex offenders: An evaluation*. Birmingham, U.K.: University of Birmingham.

*Allam, J. (1999). *Effective practice in work with sex offenders: A re-conviction study comparing treated and untreated offenders*. West Midlands Probation Service Sex Offender Unit. (Available from Jayne Allam WMPS SOU, 826 Bristol Road, Selly Oak, Birmingham, West Midlands, B29 6NA)

Andrews, D. A., Zinger, E., Hoge, R. D., Bonta, J., Gendreau, P., & Cullen, F. T. (1990). Does correctional treatment work? A clinically-relevant and psychologically informed meta-analysis. *Criminology, 28*, 369–404.

*Bakker, L., Hudson, S., Wales, D., & Riley, D. (1999). *"And there was light": An evaluation of the Kia Marama Treatment Programme for New Zealand sex offenders against children*. Unpublished report. New Zealand.

Bangert-Drowns, R. L., Wells-Parker, E., & Chevillard, I. (1997). Assessing the methodological quality of research in narrative reviews and meta-analyses. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 405–429). Washington, DC: American Psychological Association.

Barbaree, H. E. (1997). Evaluating treatment efficacy with sexual offenders: The insensitivity of recidivism studies to treatment effects. *Sexual Abuse: A Journal of Research and Treatment, 9*, 111–128.

*Barbaree, H. E., & Seto, M. C. (1998). *The ongoing follow-up of sex offenders treated at the Warkworth Sexual Behaviour Clinic*. Research Report prepared for the Correctional Service of Canada. [Warkworth]

*Barnes, J. M., & Peterson, K. D. (1997). The Kentucky Sex Offender Treatment Program. *Commonwealth of Kentucky, Justice Cabinet, Department of Corrections, Division of Mental Health*. Kentucky State Reformatory, Kentucky.

*Berlin, F. S., Hunt, W. P., Malin, M. H., Dyer, A., Lehne, G. K., & Dean, S. (1991). A five-year plus follow-up survey of criminal recidivism within a treated cohort of 406 pedophiles, 111 exhibitionists and 109 sexual aggressives: Issues and outcomes. *American Journal of Forensic Psychiatry, 12*(3), 5–28.

*Berliner, M. L., Schram, D., Miller, L. L., & Milloy, C. D. (1995). A sentencing alternative for sex offenders: A study of decision making and recidivism. *Journal of Interpersonal Violence, 10*(4), 487–502. [Washington]

*Bluglass, R. (1980). Indecent exposure in the West Midlands. In D. West (Ed.), *Sex offenders in the criminal justice system* (pp. 171–180). Cambridge: Cambridge Institute of Criminology.

*Borduin, C. M., Henggeler, S. W., Blaske, D. M., & Stein, R. J. (1990). Multisystemic treatment of adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology, 34*, 105–113.

*Borduin, C. M., Schaeffer, C. M., & Heilblum, N. (2000, May). *Multi-systemic treatment of juvenile sexual offenders: A progress report*. Paper presented at the 6th International Conference on the Treatment of Sexual Offenders, Toronto.

*Bremer, J. F. (1992). Serious juvenile sex offenders: Treatment and long-term follow-up. *Psychiatric Annals, 22*(6), 326–332.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

*Clelland, S. R., Studer, L. H., & Reddon, J. R. (1998). Follow-up report of rapists treated in a forensic psychiatric hospital. *Violence and Victims, 13*(1), 79–86. [Alberta Hospital]

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cooper, H., & Hedges, L. V. (Eds.). (1994). *The handbook of research synthesis.* New York: Sage.

*CS/RESORS. (1991). *An evaluation of community sex offender programs in the Pacific Region* (No.21803-0-A602/01-XSB). Vancouver, BC: Author.

*Davidson, P. R. (1984, January). *Behavioural treatment for incarcerated sex offenders: Post-release outcome*. Paper presented at the Sex Offender Assessment and Treatment Conference, Kingston, ON. [RTC Ontario]

*Dwyer, S. M. (1997). Treatment outcome study: Seventeen years after sexual offender treatment. *Sexual Abuse: A Journal of Research and Treatment, 9*(2), 149–160. [Dwyer]

*Earls, C. M. (1997). Étude sur la récidive. *Clinique La Macaza Bulletin d'information, 2*(1), 1–2. [La Macaza]

*Earls, C. M., Martin, I., & Bélanger, H. (October 18, 1999). *La Macaza outcome study.* Unpublished raw data. [La Macaza]

*Evaluation of sex offender rehabilitation programs in the state of Florida. (1977). Department of Health and Rehabilitative Services. Mental Health Program Office, Talahassee, FL. [Florida]

Fleiss, J. L. (1994). Measures of effect size for categorical data. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 245–260). New York: Sage.

Friendship, C., & Thornton, D. (2001). Sexual reconviction for sexual offenders discharged from prison in England and Wales. *British Journal of Criminology, 41*, 285–292.

Furby, L., Weinrott, M. R., Blackshaw, L. (1989). Sex offender recidivism: A review. *Psychological Bulletin, 105*, 3–30.

Gallagher, C. A., Wilson, D. B., Hirschfield, P., Coggeshall, M. B., & MacKenzie, D. L. (1999). A quantitative review of the effects of sex offender treatment on sexual reoffending. *Corrections Management Quarterly, 3*, 19–29.

*Gordon, A., & Packard, R. (1998). *The impact of community maintenance treatment on sex offender recidivism*. Paper presented at the 17th annual meeting of the Association for the Treatment of Sexual Abusers, Vancouver, BC. [Twin Rivers]

*Guarino-Ghezzi, S., & Kimball, L. M. (1998). Juvenile sex offenders in treatment. *Corrections Management Quarterly, 2*(1), 45–54.

*Hall, G. C. N. (1995a). The preliminary development of theory-based community treatment for sexual offenders. *Professional Psychology: Research and Practice, 26*(5), 478–483.

Hall, G. C. N. (1995b). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology, 63*, 802–809.

Hanson, R. K. (1997). How to know what works with sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 9*, 129–145.

Hanson, R. K. (2000). Treatment outcome and evaluation problems (and solutions). In D. R. Laws, S. M. Hudson, & A. Ward (Eds.), *Remaking relapse prevention with sex offenders: A sourcebook* (pp. 485–499). Thousand Oaks, CA: Sage.

Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology, 66*, 348–362.

*Hanson, R. K., & Nicholaichuk, T. (2000). A cautionary note regarding Nicholaichuk et al. (2000). *Sexual Abuse: A Journal of Research and Treatment, 12*, 289–293. [Clearwater]

*Hanson, R. K., Steffy, R. A., & Gauthier, R. (1992). *Long-term follow-up of child molesters: Risk prediction and treatment outcome* (User Report No. 1992-02). Ottawa: Corrections Branch, Ministry of the Solicitor General of Canada.

*Hanson, R. K., Steffy, R. A., & Gauthier, R. (1993). Long-term recidivism of child molesters. *Journal of Consulting and Clinical Psychology, 61,* 646–652.

Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior, 24,* 119–136.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1998). Appraisal and management of risk in sexual aggression: Implications for criminal justice policy. *Psychology, Public Policy, and Law, 4,* 73–115.

Hedges, L. V. (1994). Fixed effect models. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 285–299). New York: Sage.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* New York: Academic Press.

*Hersh, K., Baldwin, K., & Gray-Little, B. (1999, September). *Treatment completion and recidivism among incarcerated sex offenders.* Paper presented at the annual conference of the Association for the Treatment of Sexual Abusers, Orlando, FL.

*Huot, S. (1999, April). *Minnesota study of sex offender recidivism.* Unpublished data.

*Kramer, S. P. (1985). *Sex offender treatment and tracking: Problems, perspectives and outcomes in the Utah criminal justice system.* Unpublished manuscript.

*Lab, S. P., Shields, G., & Schondel, C. (1993). Research note: An evaluation of juvenile sexual offender treatment. *Crime and Delinquency, 39*(4), 543–553.

Laws, D. R. (Ed.). (1989). *Relapse prevention with sex offenders.* New York: Guilford.

*Lindsay, W. R., Marshall, I., & Neilson, C. (1998). The treatment of men with a learning disability convicted of exhibitionism. *Research in Developmental Disabilities, 19*(4), 295–316.

*Lindsay, W. R., Neilson, C. Q., & Morrison, F. (1998). The treatment of six men with a learning disability convicted of sex offences with children. *British Journal of Clinical Psychology, 37,* 83–89.

*Lindsay, W. R., & Smith, A. H. W. (1998). Response to treatment for sex offenders with intellectual disability: A comparison of men with 1- and 2-year probation sentences. *Journal of Intellectual Disability Research, 42,* 346–353.

*Looman, J., Abracen, J., & Nickolaichuk, T. P. (2000). Recidivism among treated sexual offenders and matched controls: Data from the Regional Treatment Centre (Ontario). *Journal of Interpersonal Violence, 15,* 279–290. [RTC Ontario]

Lösel, F. (1995). The efficacy of correctional treatment: A review and synthesis of meta-evaluations. In J. McGuire (Ed.), *What works: Reducing reoffending* (pp. 79–114). Chichester: Wiley.

Mander, A. M., Atrops, M. E., Barnes, A. R., & Munafo, R. (1996). *Sex offender treatment program: Initial recidivism study.* Unpublished report, Alaska Department of Corrections Offenders Programs, Alaska Justice Statistical Analysis Unit Justice Center and University of Alaska, Achorage.

*Marques, J. K. (1999). How to answer the question, "Does sex offender treatment work?" *Journal of Interpersonal Violence, 14*(4), 437–451.

*Marques, J. K., & Day, D. M. (May, 1998). *Sex offender treatment evaluation project: Progress report (May, 1998).* Sacramento, CA: California Department of Mental Health.

*Marques, J. K., Day, D. M., Nelson, C., & West, M. A. (1993). Findings and recommendations from California's experimental treatment program. In G. C. N. Hall, R. Hirschman, J. R. Graham, & M. S. Zaragoza (Eds.), *Sexual aggression: Issues in etiology, assessment and treatment* (pp. 197–214). Washington, DC: Taylor and Francis.

*Marques, J. K., Day, D. M., Nelson, C., & West, M. A. (1994). Effects of cognitive–behavioral treatment on sex offenders recidivism: Preliminary results of a longitudinal study. *Criminal Justice and Behavior, 21,* 28–54.

*Marshall, W. L., & Barbaree, H. E. (1988). The long-term evaluation of a behavioral treatment program for child molesters. *Behaviour Research and Therapy, 26*(6), 499–511.

*Marshall, W. L., Eccles, A., & Barbaree, H. E. (1991). The treatment of exhibitionists: A focus on sexual deviance versus cognitive and relationship features. *Behaviour Research and Therapy, 29,* 129–135.

*Martin, I. (1998). *Efficacité d'un programme cognitif-behavioural institutionnel pour délinquants sexuels.* Unpublished doctoral dissertation, Department of Psychology, University of Montreal. [La Macaza]

*McGrath, R. J., Hoke, S. E., & Vojtisek, J. E. (1998). Cognitive-behavioral treatment of sex offenders. *Criminal Justice and Behavior, 25*(2), 203–225.

*McGuire, T. J. (2000). Correctional institution based sex offender treatment: A lapse behavior study. *Behavioural Sciences and the Law, 18,* 57–71.

*Meyer, L. C., & Romero, J. (1980). *A ten year follow-up of sex offender recidivism.* Pennsylvania, PA: Joseph J. Peters Institute. [JJ Peters]

*Miner, M., & Dwyer, S. M. (1995). Analysis of dropouts from outpatient sex offender treatment. *Journal of Psychology and Human Sexuality, 7*(3), 77–93. [Dwyer]

*The Missouri Sexual Offender Program. (1989). *Updated figures for Missouri sex offender study.* Department of Corrections, Division of Classification and Treatment, Jefferson City, MO. [Missouri]

*Mulloy, R., & Smiley, W. C. (1996). *Recidivism and treated sex offenders.* Paper presented at the International Congress of Psychologists, Montreal, QC. [RHC Pacific]

*Nicholaichuk, T. P. (1996). Sex offender treatment priority: An illustration of the risk/need principle. *Forum on Corrections Research, 8*(2), 30–32. [Saskatchewan]

Nicholaichuk, T., Gordon, A., Andre, G., & Gu, D. (1995). *Outcome of the Clearwater Sex Offender Treatment Program: A matched comparison between treated and untreated offenders.* Unpublished manuscript, Regional Psychiatric Centre, Saskatoon, Saskatchewan, Canada.

*Nicholaichuk, T., Gordon, A., Gu, D., & Wong, S. (2000). Outcome of an institutional sexual offender treatment program: A comparison between treated and matched untreated offenders. *Sexual Abuse: A Journal of Research and Treatment, 12,* 139–153. [Clearwater]

*Nutbrown, V., & Stasiak, E. (1987). Research monograph: A retrospective analysis of O.C.I. cost effectiveness 1977–1981. *Ontario Correctional Institute, 2,* 1–16.

*Pellerin, B., Proulx, J., Ouimet, M., Paradis, Y., McKibben, A., & Aubut, J. (1996). Étude de la recidive post-traitement chez des aggresseurs sexuels judiciarises. *Criminologie, 29,* 85–108. [Pinel]

*Perkins, D. (1987). A psychological treatment programme for sexual offenders. In B. J. McGurk, D. M. Thornton, & M. Williams (Eds.), *Applying Psychology to Treatment: Theory & Practice* (pp. 192–217). Her Majesty's Stationary Office, London.

*Pierson, T. A. (1988). *The Missouri Sexual Offender Program: Inmate Characteristics and Recidivism Analysis, 1984–1985 releases.* Department of Corrections, Division of Classification and Treatment, Jefferson City, MO. [Missouri]

*Proctor, E. (1996). A five-year outcome evaluation of a community-based treatment program for convicted sexual offenders run by the probation service. *Journal of Sexual Aggression, 2,* 3–16.

*Proulx, J., Ouimet, M., Pellerin, B., Paradis, Y., McKibben, A., & Aubut, J. (1998). *Posttreatment recidivism rates in sexual aggressors: A comparison between dropout and nondropout subjects* (No. 98-22). Les cahiers de l'ecole de criminologie, University of Montreal, Montreal, Quebec. [Pinel]

*Quinsey, V. L., Khanna, A., & Malcolm, P. B. (1998). A retrospective evaluation of the Regional Treatment Centre sex offender treatment program. *Journal of Interpersonal Violence, 13,* 621–644. [RTC Ontario]

*Rattenbury, F. R. (1986). *The outcomes of hospitalized and incarcerated sex offenders: A study of offender types, recidivism rates, and identifying characteristics of the repeat offender.* Unpublished doctoral dissertation, School of Loyola University of Chicago, IL.

Redondo, A., Sánchez-Meca, J., & Garrido, V. (1999). The influence of treatment programmes on the recidivism of juvenile and adult offenders: An European meta-analytic review. *Psychology, Crime and Law, 5,* 251–278.

*Report on treatment programs for sex offenders: Summary and recommendations. (1984). In conjunction with the Governor's Task Force on Sex Offenders and Their Victims, Florida Mental Health Institute, Tampa, FL. [Florida]

*Rice, M. E., Quinsey, V. L., & Harris, G. T. (1991). Sexual recidivism among child molesters released from a maximum security psychiatric institution. *Journal of Consulting and Clinical Psychology, 59,* 381–386.

*Robinson, D. (1995). *The impact of cognitive skills training on post-release recidivism among Canadian federal offenders* (No. R-41). Ottawa, ON, Correctional Service Canada, Correctional Research and Development.

*Romero, J. J., & Williams, L. M. (1983). Group psychotherapy and intensive probation supervision with sex offenders. *Federal Probation, 47,* 36–42. [JJ Peters]

*Seto, M. C., & Barbaree, H. E. (1999). Psychopathy, treatment behavior, and sex offender recidivism. *Journal of Interpersonal Violence, 14,* 1235–1248. [Warkworth]

*Sex offender programs: Program manuals, descriptions, evaluations, & research. (1998). Regional Health Centre (Pacific). [RHC Pacific]

*Sex offenses in Washington State: 1998 update. (1998). Washington State Institute for Public Policy, Olympia, WA. [Washington]

*Smiley, W. C., & Mulloy, R. (1995). Treatment of sex offenders: Does intervention work? Unpublished Manuscript. [RHC Pacific]

*Song, L., & Lieb, R. (1995). Washington state sex offenders: Overview of recidivism studies. Olympia, WA: Washington State Institute for Public Policy. [Twin Rivers; Washington]

*Studer, L. H., & Reddon, J. R. (1998). Treatment may change risk prediction for sexual offenders. Sexual Abuse: A Journal of Research and Treatment, 10, 181. [Alberta Hospital]

*Studer, L. H., Reddon, J. R., Roper, V., & Estrada, L. (1994). Sexual offender recidivism: An evaluation of the Pheonix program. Alberta Hospital, Edmonton, AB. [Alberta Hospital]

Swenson, C. C., Henggeler, S. W., & Schoenwald, S. K. (2001). Family-based treatments. In C. R. Hollin (Ed.), Handbook of offender assessment and treatment (pp. 205–220). Chichester, U.K.: Wiley.

United States General Accounting Office. (1996). Sex offender treatment: Research results inconclusive about what works to reduce recidivism. Washington, DC: Author.

*Walker, D. W. (2000). The treatment of adult male child molesters through group family intervention. In A. P. Colmenárez & R. H. Serrano (Eds.), Violencia (pp. 191–197). Caracas, Venezuela: Universidad Central de Venezuela.

Wierzbicki, M., & Pekarik, G. (1993). A meta-analysis of psychotherapy drop-out. Professional Psychology: Research and Practice, 24, 190–195.

Wille, R., & Beier, K. M. (1989). Castration in Germany. Annals of Sex Research, 2, 103–133.

*Worling, J. R., & Curwen, T. (1998). The adolescent sexual offender project: A 10-year follow-up study. SAFE-T Program (Sexual Abuse: Family Education & Treatment) Thistletown Regional Centre for Children & Adolescents & Probation and Community Services, Ontario Ministry of Community and Social Services: Toronto, ON.

The effectiveness of treatment for sexual offenders:
A comprehensive meta-analysis

FRIEDRICH LÖSEL and MARTIN SCHMUCKER*
University of Erlangen, Nuremberg, Germany
*corresponding author: E-mail: mnschmuc@phil.unirlangen.de

Abstract. The article reports a meta-analysis on controlled outcome evaluations of sexual offender treatment. From 2,039 documents published in five languages, 69 studies containing 80 independent comparisons between treated and untreated offenders fulfilled stepwise eligibility criteria (total N = 22,181). Despite a wide range of positive and negative effect sizes, the majority confirmed the benefits of treatment. Treated offenders showed 6 percentage points or 37% less sexual recidivism than controls. Effects for violent and general recidivism were in a similar range. Organic treatments (surgical castration and hormonal medication) showed larger effects than psychosocial interventions. However, this difference was partially confounded with methodological and offender variables. Among psychological programs, cognitive–behavioral approaches revealed the most robust effect. Nonbehavioral treatments did not demonstrate a significant impact. There was no outcome difference between randomized and other designs, however, group equivalence was associated with slightly larger effects. Various other moderators had a stronger impact on effect size (e.g., small sample size, quality of outcome reporting, program completion vs. dropout, age homogeneity, outpatient treatment, and authors' affiliation with the program). More differentiated, high-quality evaluations are needed to clarify: What works for whom under which circumstances?

Key words: evaluation, meta-analysis, recidivism, sex offender treatment, treatment efficacy

Abbreviations: CG = comparison group; TG = treatment group; ES = effect size; OR = odds ratio; LOR = logged odds ratio; $CI_{95\%}$ = 95% confidence interval

Treatment for sexual offenders is a very important topic of criminal policy. Media reports on serious cases of sexually motivated murder, rape, and child abuse have made people particularly concerned about this area of crime. In various countries, policymakers have reacted by increasing measures of both punishment and treatment. Because most incarcerated sexual offenders return to the community, effective treatment is a cornerstone for preventing future offenses. However, the empirical basis of sex offender treatment is less solid than such a cornerstone should be. Although recent overviews suggest a moderately positive effect (e.g., Alexander 1999; Aos et al. 2001; Gallagher et al. 2000; Hanson et al. 2002; Lösel 2000; Polizzi et al. 1999), methodological problems, inconsistent results, and a lack of high quality studies question how far we know what works for sex offenders (e.g., Lösel 2000; Marques et al. 2005; Quinsey et al. 1993; Rice and Harris 2003; White et al. 1998).

Indeed, there is much less well-controlled research on the evaluation of programs for sex offenders than in the field of general offender treatment (Lipsey and Wilson 1998; Lösel 2001a; McGuire 2002). Clear messages on the efficacy of programs for

sex offenders are complicated by various problems that relate to issues of design and treatment.

Design issues.   Serious offenders relatively often receive at least some kind of treatment. This makes it difficult to form randomized untreated control groups. The relatively low baserate of official sexual reoffending (e.g., Hanson and Bussière 1998) means that large samples are needed to reveal a significant effect. However, for economic reasons, evaluations of sex offender treatment often contain rather small samples. This enhances the risk of nonequivalent groups, even in randomized designs (Marques et al. 2005). Furthermore, selective dropout and other problems make it difficult not only to implement but also to maintain internal validity (e.g., Barbaree 1997; Marshall and McGuire 2003). In addition to the general problems of undetected crime, the registration of sexual offenses is influenced by the intimate nature of this kind of crime. This limits the value of official data.

Treatment issues.   The analysis of sexual reoffending requires relatively long follow-up periods (Prentky et al. 1997), so that evaluations may address old programs that do not represent the current 'state of the art.' Furthermore, the treatment of sex offenders embraces a broad range of medical and psychosocial programs, and sex offenders are a rather heterogeneous group. As a result, adequate evaluations and replications need to be rather differentiated. Moreover, sex offenders often do not receive a single program but individualized 'packages' (e.g., psychological and pharmacological treatment; additional programs for alcohol dependence). Disentangling the impact of specific modules would require relatively complex evaluations.

These and other problems make systematic research integrations more difficult than in those areas in which we can draw on a large number of methodologically sound evaluations. Accordingly, an early meta-analysis from Furby et al. (1989) found few well-designed studies and concluded that there is "as yet no evidence that clinical treatment reduces rates of sex reoffenses" (p. 27). Nine years later, a review for the Cochrane Collaboration came to a similar conclusion (White et al. 1998). Restricting their analysis to randomized evaluations, the authors ended up with only three studies. Other recent meta-analyses included quasi-experimental studies as well. This is in line with the notion that methodological quality is not a unidimensional all-or-none category, and that practice can also learn from good but suboptimal studies (Cronbach et al. 1980; Lösel and Köferl 1989; Shadish et al. 2002).

Most reviews report a lower average rate of sexual recidivism in treated groups than in control groups. For example, Hall (1995) integrated 12 controlled studies. The average rate of sexual recidivism was 19% in treated groups and 27% in controls (mean effect size (ES): $d = 0.24$). Hormonal and cognitive–behavioral treatments were more successful than other programs. However, such differential effects based on a small number of studies and a single evaluation of castration had a strong impact. Alexander (1999) integrated 79 studies on psychosocial sex offender treatment. The mean difference in recidivism was 5 percentage points in favor of treatment ($d = 0.12$; Lösel and Schmucker 2003). However, the majority

of studies contained no control group (Level 1 on the Maryland Scale of Methodological Rigor; Sherman et al. 1997). The same holds for a meta-analysis of 20 studies on medical and psychological treatment from German-speaking countries. This revealed a nonsignificant effect of $d = 0.08$ (Lösel 2000). Gallagher et al. (2000) analyzed 23 studies of relatively good quality. Treated groups showed 10 percentage points less sexual recidivism than controls and the overall effect size was relatively large ($d = 0.47$). In the most comprehensive recent meta-analysis, Hanson et al. (2002) integrated 43 studies on psychological treatment. The average sexual offense recidivism was 12.3% for treatment groups and 16.8% for comparison groups (mean ES: $d = 0.13$; Lösel and Schmucker 2003). Current, mainly cognitive–behavioral treatment revealed better outcomes whereas older programs appeared to have little effect.

Overall, the last decade has shown a strong increase and more positive outcomes in evaluations of sex offender treatment. However, research syntheses vary in effect size, type of treatment included, prevailing design quality, categorization of programs, treatment settings, and meta-analytical techniques (Lösel and Schmucker 2003). Nearly all are restricted to English-language studies that mainly come from North America. Against this background, an updated, comprehensive, independent, and international review seems worthwhile. This is why the Campbell Collaboration Group on Crime and Justice (Farrington and Petrosino 2001) has included a systematic review on sex offender treatment in its portfolio. The present study is a first report of our work on this synthesis. It contains a meta-analysis of studies on both psychological and biological treatment that were published up to 2003 in English and various other languages.

Method

Criteria for inclusion of studies

Primary studies had to have the following characteristics to be eligible:

1. *Study of sexual offenders.* Participants had to have been convicted of a sexual offense or to have committed acts of illegal sexual behavior that would have lead to a conviction if officially prosecuted.
2. *Evaluation of treatment.* No restrictions were made on the kind of intervention applied as long as it aimed to reduce recidivism (i.e., psychosocial as well as organic treatment modes such as hormonal medication or surgical castration were included). However, interventions had to incorporate therapeutic measures; purely deterrent or punishing approaches were not included. Treatment did not have to be specifically tailored for sexual offenders. More general measures of offender treatment were included if the study addressed at least a subgroup of sexual offenders and reported separate results for sexual offenders in both the treated and control groups.

3. *Measure of recidivism as outcome.* Recidivist behavior had to be included as a dependent variable. We followed a broad definition of recidivism and included outcomes ranging from incarceration to lapse behavior. In contrast, primary studies focusing exclusively on changes in measures of personality or hormone levels, clinical ratings of improvement, and the like were not included.

4. *Control group design.* The study had to report the same recidivism outcome for a comparison group not receiving the same treatment. This could either be an untreated control group or a group of offenders receiving treatment 'as usual' or another kind of treatment that differed from the evaluated program in content, intensity and specificity. Studies using only treatment dropouts as a control group were not eligible.

5. *Sample size.* Studies had to contain a minimum total sample size of 10 persons with at least five offenders in each group. In very small samples, even minimal differences in absolute numbers of recidivism between groups may lead to extreme effect sizes, thus making results very heterogeneous while not being very reliable.

6. *Sufficient data for effect size computation.* Studies had to report outcomes in a way permitting the calculation of effect size estimates.

7. *Country of origin.* No restrictions were made as to where studies were conducted. For economic reasons, we restricted our analysis to studies reported in English, German, French, Dutch, or Swedish.

8. *Time of publication.* There were no restrictions regarding the time of publication. All studies reported up to June 2003 were eligible.

Literature search

Searches were designed to tap published as well as unpublished evaluations. A wide range of sources was used to identify relevant studies. First, a basic study pool was compiled from the reference sections of previous reviews. Second, already identified primary studies were analyzed for further references. Third, 14 electronic databases were analyzed (e.g., PsycInfo, MedLine, ERIC, Cochrane Library, Social Services Abstracts, NCJRS abstract and full-text databases, Dissertation Abstracts International, and UK National Health Service National Register). Fourth, we hand searched journals pertaining to the topic (e.g., Sexual Abuse, Journal of Sexual Aggression). To identify more unpublished work, we contacted researchers in the field of sexual offender treatment and asked if they knew or had personally conducted further evaluations. Finally, an Internet search was conducted. Because the Internet constitutes a vast pool of rather loosely connected information, it cannot be searched in total. We visited the Internet sites of pertinent institutions and Departments of Corrections and searched them and their links for relevant material.

A total of 2,039 citations were identified in this way. These were scanned by their title first. If not obviously off-theme, we then examined the abstracts. If these

indicated potential eligibility or if no abstract was available, we used the full reports to finally evaluate eligibility. Twenty-one (1%) of the reports could not be obtained (of which 18 also had no available abstract). A total of 586 citations were clearly off-topic; 242 referred to offenders in general; and 641 did not evaluate a program, that is, they either simply described a treatment or reviewed sex offender treatment more generally. This left 549 citations referring to primary studies evaluating sexual offender treatment. Of these, 189 were excluded because they did not report recidivism outcome, 236 did not employ an adequate comparison group, 56 duplicated evaluations already included, and two did not lend themselves to effect size calculation. In the end, 66 reports met the specified inclusion criteria. Some contained more than one eligible study. In such cases, we referred to the individual studies as the unit of analysis. Some primary studies presented results for different subgroups (e.g., offense types). To allow for maximum differentiation while adhering to the principle of independency between effect sizes, we chose these subgroups as units of analysis. However, the characteristics of the subgroups had to be reported in just as much detail as the total sample. Following these rules, we formed a database of 80 comparisons from 69 studies.

Coding of study characteristics

Coding followed a detailed manual. A selection of variables is presented in the Results section (see Table 1). The manual covered general characteristics of the publication (e.g., year, country, type of publication), sample description (e.g., sample size, types of offender, age, voluntariness of treatment participation), methodological features (e.g., study design, type of control group, follow-up interval, source of recidivism data), and characteristics of the treatment (e.g., mode, setting, integrity, and format of treatment). Various programs for sex offenders are not restricted to a certain therapeutic paradigm but combine strategies from different 'schools' in an eclectic manner. Thus, as well as categorizing the basic therapeutic approach, we also rated the degree to which the different treatment elements were applied in an individual intervention on separate 4-point scales.

We evaluated the overall methodological quality of the individual study with an integrative rating scheme (see, for various options, Farrington 2003). We adapted the Maryland Scale of Scientific Rigor (Sherman et al. 1997) for our purposes. This is a 5-point scale integrating methodological features related to the validity of a causal interpretation of treatment effects. The highest level (5) is reserved for uncompromised random designs. Level 4 covers studies applying procedures to ensure group equivalence (e.g., individual matching, statistical control) or slightly compromised random designs. Designs based on incidental assignment are on Level 3 if group equivalence can be assumed (e.g., demonstrated equivalence on relevant variables). Studies incorporating a nonequivalent control group correspond to Level 2. Finally, Level 1 is reserved for uncontrolled studies that were not eligible for our meta-analysis. In its original form, the Maryland Scale also covers sample size and adequacy of statistical testing. However, because we integrated

studies by means of weighted effect sizes, these aspects were not so relevant and we dropped them in our rating.

Following Lösel and Köferl (1989), we included ratings on the completeness and accuracy of information reported. On a 4-point scale, these measured threats to 'descriptive validity' in the domains of treatment concept, treatment implementation, assessment of treatment goals, and methods used.

Studies were coded by the second author. A subsample of 10 studies was additionally coded by an experienced rater. The average interrater agreement was 91%. Core variables such as treatment type or design showed full agreement. No variable fell below 60%.

Computation of effect size

As an effect size measure, we used odds ratios (OR). This is widely recommended for dichotomous data (Fleiss 1994; Lipsey and Wilson 2001). Results in primary studies on offender treatment are usually reported as simple recidivism rates (P) or as the absolute number of successes and failures in the treatment group ($TG_{Success}$, $TG_{Failure}$) and the comparison group ($CG_{Success}$, $CG_{Failure}$), respectively. In such cases, we applied the following formulas for effect size computation:

$$OR = \frac{P_{CG} \times (1 - P_{TG})}{P_{TG} \times (1 - P_{CG})} \quad \text{and} \quad OR = \frac{CG_{Failure} \times TG_{Success}}{TG_{Failure} \times CG_{Success}}$$

If any of these frequencies equaled zero, 0.5 was added to each frequency. Some studies reported more sophisticated statistical analyses that controlled for differences between treatment and control groups. In such cases, we used these results instead of the simple recidivism rates. In logistic regression, the coefficients equal the natural log of the odds ratio (LOR), and as an exponent to e this equals the odds ratio (see Fleiss 1994). The result for the treatment variable could thus be transferred directly. In Cox regression, results are reported in the form of a rate ratio, which is similar but not identical to the odds ratio. We used the rate ratio (RR) to estimate a recidivism rate for the control group corrected for the group differences considered in the Cox regression model ($P_{CG} = RR \times P_{TG}$ or $P_{CG} = RR/P_{TG}$, depending on the coding of the treatment variable in the primary study). We then calculated the odds ratio with the above formula. Few studies reported other test statistics that could not be transformed readily into odds ratios. In these cases, we used standard procedures to calculate Cohen's d (see Lipsey and Wilson 2001) and then converted these into odds ratios using $LOR = \frac{\pi}{\sqrt{3}} \times d$ (Hasselblad and Hedges 1995, Formula 4, re-arranged) and $OR = e^{LOR}$.

If a study contained multiple dependent (sub-)samples, we used the comparison with the highest internal validity (e.g., if a study compared recidivism rates for the total sample of treated/untreated participants and additionally matched a subsample of these groups on relevant characteristics, we would use the latter comparison). Studies often reported multiple outcome variables. Different domains of recidivist behavior (i.e., sexual, violent, or general recidivism) were always analyzed separately. If a

study used different indicators of failure (i.e., charge, arrest, conviction, or lapse behavior), effect sizes were coded separately and then averaged to a single effect size.

Some studies reported separate results for different offender types or risk groups, but did not meet criteria for independent comparisons as defined above. Here, we calculated effect sizes separately for the subgroups and used the weighted average to obtain a study effect size (see Fleiss 1994).

Whenever possible, participants who dropped out of treatment were included in the treatment group ('intent to treat'). However, we evaluated the effects of treatment dropout by additionally contrasting them with both treated and un-treated groups.

Integration and statistical analyses

Statistical analyses were conducted on the natural log of the odds ratio (see Fleiss 1994; Lipsey and Wilson 2001). To integrate effect sizes, we applied the weighting procedures based on the standard error of individual effect sizes (Hedges and Olkin 1985). Because of heterogeneous effect size distributions (according to the Q test of homogeneity; Hedges and Olkin 1985), we applied a random effects model. Moderator analyses were carried out under the assumption of a mixed effects model (see, also, Lipsey and Wilson 2001; Wilson 2001).

Results

Descriptive characteristics

Table 1 gives an overview of some characteristics of the studies/comparisons included for analysis. Most studies came from North America. Approximately one-third contained unpublished data. Most were relatively recent. Nearly three quarters of the studies have been published since 1990. However, as the treatment section shows, the actual program implementation started much earlier.

Nearly one-half of the comparisons addressed cognitive–behavioral programs. Due to basic similarities, we also subsumed two studies of multisystemic treatment under this category. Fourteen comparisons addressed physical therapy, eight of which dealt with surgical castration. In more than one-half of the studies, authors were affiliated with the evaluated treatment. Most treatments were specifically designed for sex offenders. However, it was rarely possible to rate whether treatment was implemented reliably. Only one-quarter of the comparisons provided a documentation of adequate program integrity. Approximately one-half of the interventions took place in an institutional setting. Although a group format was most frequent, nearly one-half of the programs included at least some individual-ized treatment. An explicit extension of treatment through specific aftercare services was reported for only 15 comparisons.

Table 1. Descriptive characteristics of the 80 comparisons included in the meta-analysis.

Coding variable and categories	Frequency	Percentage
General study characteristics[a]		
Publication year		
Before 1980	7	10.1
1980s	11	15.9
1990s	29	42.0
Since 2000	22	31.9
Country		
USA	31	44.9
Canada	17	24.6
Great Britain	8	11.6
German-speaking countries	8	11.6
Other	5	7.2
Publication type		
Journal article	34	49.3
Book, chapter	10	14.5
Unpublished	25	36.2
Treatment characteristics		
Time of treatment implementation		
Before 1970	14	17.5
1970s	17	21.3
1980s	30	37.5
1990s	19	23.8
Mode of treatment		
Cognitive-behavioral	37	46.3
Classical behavioral	7	17.5
Insight-oriented	7	17.5
Therapeutic community	10	12.5
Other psychosocial, unclear	5	6.3
Hormonal medication	6	7.5
Surgical castration	8	10.0
Author affiliation to treatment program		
Yes	42	52.5
No	31	38.8
Unclear	7	8.8
Sex-offender-specific treatment		
Yes	64	80.0
No	9	11.3
No information available	7	8.8
Integrity of treatment implementation[b]		
Acceptable	18	25.0
Problematic	5	6.9
No information available	49	68.1
Setting of treatment		
Prison	25	31.3
Hospital	14	17.5
Outpatient	29	36.3
Mixed	10	12.5
No information available	2	2.5

Table 1. Continued.

Coding variable and categories	Frequency	Percentage
Format of treatment[b]		
Only group treatment	18	22.5
Mainly group treatment	17	21.3
Mixed	10	12.5
Mainly individual treatment	8	10.0
Only individual treatment	7	8.8
No information available	12	15.0
Aftercare		
Obligatory	9	11.3
Optional	6	7.5
Not offered, not reported	65	81.3
Offender characteristics		
Age group		
Adults	45	56.3
Adolescents	7	17.5
Mixed	8	10.0
Unclear	20	25.0
Homogeneity of age		
High	7	17.5
Medium	23	28.8
Low	23	28.8
Unclear	27	33.8
Offense type[c]		
Rape	44	55.0
Child molestation	59	73.7
Incest offenses	38	47.5
Exhibitionism	24	30.0
Other hands-on offenses, not specified	5	6.3
Other hands-off offenses, not specified	5	6.3
Not specified	20	25.0
Treatment participation		
Voluntary	37	46.3
Nonvoluntary	16	20.0
Mixed	8	10.0
Unclear	19	23.8
Methodological characteristics		
Sample size		
Up to 50	25	31.3
51–100	12	15.0
101–200	18	22.5
201–500	14	17.5
More than 500	11	13.3
Maryland scale		
Level 2 (nonequivalent)	48	60.0
Level 3 (equivalence assumed)	19	23.8
Level 4 (matching, statistical control)	7	8.8
Level 5 (randomization)	6	7.5

Table 1. Continued.

Coding variable and categories	Frequency	Percentage
Initial group equivalence		
Yes	20	25.0
No, TG at higher risk	14	17.5
No, CG at higher risk	4	5.0
No, direction unclear	8	10.0
No, but controlled for statistically	5	6.3
No information available	29	36.3
Control group formation		
Treatment volunteers	13	16.3
No treatment available	14	17.5
No treatment order	15	18.8
Treatment refused	19	23.8
Other	10	12.5
Unclear	9	11.3
Type of reoffense[c]		
Sexual	74	92.5
Violent	20	25.0
Any	49	61.3
Follow-up period (months)		
12–24	14	17.5
25–36	12	15.0
37–60	23	28.8
61–84	12	15.0
>84	19	23.8
Source of recidivism data		
Criminal records only	64	80.0
Self-report	6	7.5
Not indicated	10	12.5
Definition of recidivism		
Arrest	19	23.8
Conviction	24	30.0
Charge	15	18.8
Lapse behavior	3	3.8
Multiple outcomes	6	7.5
Not indicated	13	16.3

[a]$n = 69$.
[b]$k = 72$ (no reasonable categorization for surgical castration possible).
[c]Individual comparisons may cover multiple categories.

Seven comparisons referred to programs that targeted exclusively juvenile sexual offenders. Only these were very homogeneous in terms of age. Most treatment programs combined individuals with different types of sex offense. Child molestation was most frequent, followed by rape. However, no program referred exclusively to rapists. Nine programs addressed child molesters only, and four addressed exhibitionists only. Most frequently, treated offenders had participated voluntarily. Thirty percent of the comparisons referred to offenders who were at least partially obliged to attend treatment.

Sample sizes (TG + CG) varied from 15 to 2,557 (Md = 118). Roughly one-third of the comparisons contained less than 50 offenders. Only seven comparisons were based on a randomized design. One of these was compromised, so that only six could be coded on Level 5 of the Maryland Scale. Sixty percent of the comparisons were on Maryland Scale Level 2, that is, group equivalence could not be assumed. When group differences actually were tested and reported, the TG was more often at higher risk than the CG. However, for 29 comparisons (including all randomized trials), no information was available on group differences. In nearly one-quarter of the comparisons the CG consisted of treatment refusers.

Nearly all studies used a specific measure of sexual recidivism. Recidivism was recorded after an average follow-up period of more than 5 years (TG: M = 63.54 months, SD = 42.09; CG: M = 62.41, SD = 42.37). It was mainly based on entries in official criminal records. Few studies additionally used information from the offenders themselves. The most common definition of failure was reconviction, followed by rearrest and new charges. In three comparisons the criterion was defined rather loosely as 'inappropriate' or 'lapse' behavior. In six comparisons, outcomes were reported separately for different definitions.

Total effects

The 74 comparisons reporting data on sexual recidivism revealed an average recidivism rate of 12% for treated groups and 24% for comparison groups (unweighted average). This is a 50% reduction. However, when we calculated the recidivism rates for treated and comparison participants taking the respective sizes of TG and CG in the 74 comparisons into account (i.e., when we calculated an n-weighted average for treated and comparison groups), the difference in recidivism rates vanished completely (11% each for treated and comparison participants). These conflicting results were due to great differences in the size of TG and CG in some primary studies. Studies with very large control groups and comparatively small treated groups and an overall low recidivism baserate reduced the n-weighted average of the CG recidivism rate considerably but not the TG average. Therefore, one should not draw conclusions from these averages. Effect-size integration avoids the problem of different TG/CG sizes, because the recidivism rates are first evaluated on the level of the individual comparisons and only then integrated. The mean effect size can then be used to estimate mean CG recidivism to illustrate results. Figure 1 gives an overview of the effect sizes (logged OR) of the 74 comparisons. It shows that the majority of effects were positive (k = 53). Converted to OR, the effects ranged from a minimum of 0.17 to a maximum of 33.33.

We integrated the individual effect sizes according to the random model. Results are shown in Table 2. The mean OR of 1.70 for sexual recidivism was highly significant (z = 4.96, P < 0.001). The absolute difference in sexual recidivism between TG and CG was 6.4 percentage points. This is a 37% reduction
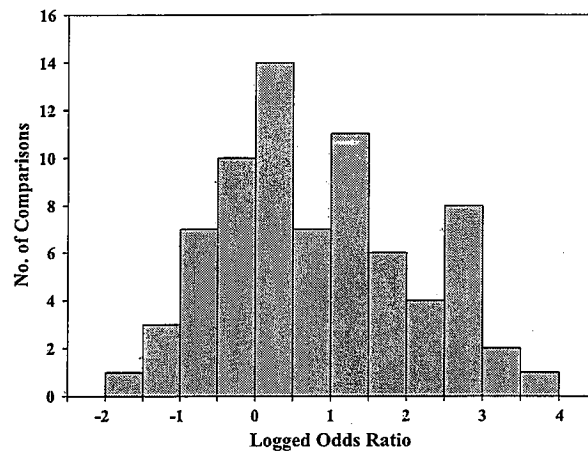
Figure 1. Distribution of effect sizes (logged odds ratios) for sexual recidivism ($k = 74$).

from the base rate of the CG. Mean effects were similar for other areas of offending. For violent recidivism, the mean OR was 1.90 ($z = 5.36$, $P < 0.001$). The average recidivism rate for treated offenders was 5.2 percentage points lower than that for untreated offenders (44% reduction). For any recidivism, the mean effect size was $OR = 1.67$ ($z = 4.52$, $P < 0.001$). The corresponding rate of general recidivism for treated offenders was 11.1 percentage points lower than for untreated offenders (31% reduction).

Except for violent recidivism, the effect size distributions showed considerable heterogeneity, Q (73) = 237.14, $P < 0.001$ for sexual recidivism and Q (48) = 159.80, $P < 0.001$ for general recidivism. Therefore, we analyzed variables that may have influenced the treatment effects. We restricted our moderator analyses to sexual recidivism as an outcome because this is the main area of interest and provides the largest database.

Table 2. Total mean effects of treatment for different areas of recidivism.

Outcome	k	OR	$CI_{95\%}$	Q	Recidivism (%)	
					TG^a	CG^b
Sexual recidivism	74	1.70***	1.35–2.13	237.14***	11.1	17.5
Violent recidivism	20	1.90***	1.49–2.33	19.68	6.6	11.8
Any recidivism	49	1.67***	1.33–2.08	159.80***	22.4	32.5

k = Number of comparisons, OR = mean odds ratio, $CI_{95\%}$ = 95% confidence interval, Q = test of homogeneity (χ^2, $df = k - 1$), TG = treated group; CG = comparison group.
***$P < 0.001$.
[a]n-Weighted average.
[b]Estimated recidivism rate.

Moderator analyses

For reasons of space, we only report a selection of moderator analyses. Table 3 reports the results on variables that revealed significant differences or are particularly relevant for treatment practice. In all analyses, we have to take into account that moderating effects may be confounded with the influence of other variables.

Treatment characteristics.   The various treatment approaches differed considerably in effect size. In total, physical treatment had higher effects (OR = 7.37, $CI_{95\%}$: 4.14–13.11, z = 6.80, P < 0.001) than did nonphysical (psychosocial) interventions (OR = 1.32, $CI_{95\%}$: 1.07–1.62, z = 2.60, P = 0.01), Q (1, k = 66) = 30.47, P < 0.001. This was particularly due to the very large mean ES for surgical castration. However, hormonal treatment also showed a higher effect than any of the psychosocial measures. Of these, only cognitive–behavioral treatments and classic behavior therapy had a significant impact on sexual recidivism. With OR close to 1, the other approaches did not influence recidivism significantly. A regression analysis entering the more differentiated 4-point scale ratings on the use of individual treatment elements confirmed these findings. It showed significant standardized regression weights (β) for a cognitive orientation (0.36), behavioral conditioning techniques (0.26), and hormonal medication (0.26); all P < 0.05.

In the further moderator analyses, we excluded the studies on surgical castration for the following reasons: (a) This approach differs strongly from the others. (b) It is currently rarely used in practice. (c) The effect sizes for castration were extremely large and homogeneous, Q (7) = 1.32, P = 0.99, and thus would have unduly biased the results.

For the remaining 66 comparisons, the mean effect remained significant (OR = 1.38, $CI_{95\%}$: 1.13–1.69, z = 3.16, P < 0.01). A test of homogeneity still indicated a highly heterogeneous effect size distribution, Q (65) = 163.92, P < 0.001, justifying further moderator analyses.

As Table 3 shows, the decade in which the program was implemented related significantly to effect size. However, there was no linear relationship. This was also confirmed in a correlation analysis (r = 0.16, P = 0.15). Thus, more 'modern' programs did not generally prove to be particularly successful. The year of publication as another indicator of recency showed a similar picture. Here, the differences were even less pronounced (r = 0.08, P = 0.51). When the study authors had been involved in the treatment program, there was a larger effect. When there was no such affiliation, the respective OR was close to 1.

Only programs designed specifically for sex offenders had a significant effect. The few others even showed a negative outcome. Although the setting variable revealed no significant difference, there was a strong tendency for relatively larger effects in outpatient treatment and smaller effects in institutions. Mixed settings had an intermediate ES. A linear order from institutional to outpatient treatment showed a significant correlation (r = 0.27, P = 0.02).

Whether the treatment was delivered in an individual or a group format did not result in significant outcome differences. However, in this category, we must assume

Table 3. Moderator analyses.

Variables	k	Q_{bet}	OR	CI$_{95\%}$
				Lower–upper
Treatment characteristics				
Treatment approach		36.02***		
Cognitive–behavioral	35		1.45**	1.12–1.86
Classic behavioral	7		2.19**	1.22–3.92
Insight oriented	5		0.98	0.51–1.89
Therapeutic community	8		0.86	0.54–1.35
Other psychosocial, unclear	5		0.94	0.53–1.65
Hormonal medication	6		3.08**	1.40–6.79
Surgical castration	8		15.34***	7.34–32.05
Time of treatment implementation		7.37**		
Before 1970	5		0.56*	0.32–0.98
1970s	14		2.03***	1.34–3.09
1980s	30		1.38**	1.08–1.77
1990s	17		1.27	0.86–1.87
Specific treatment sex offenders		4.70*		
Yes	56		1.56***	1.27–1.93
No	5		0.76	0.41–1.41
Setting of treatment		5.10		
Prison	21		1.16	0.84–1.60
Hospital	8		1.10	0.62–1.94
Outpatient	27		1.93***	1.35–2.77
Mixed	10		1.37	0.78–2.41
Format of treatment		6.74		
Only group treatment	17		1.12	0.76–1.66
Mainly group treatment	17		1.57*	1.02–2.42
Mixed	8		2.45*	1.36–4.40
Mainly individual treatment	8		1.40	0.77–2.53
Only individual treatment	6		2.88*	1.14–7.24
Author affiliation to treatment program		10.95***		
Yes	32		1.92***	1.44–2.56
No	30		0.99	0.76–1.29
Offender characteristics				
Age group		1.19		
Adolescents only	7		2.35*	1.01–5.43
Adults only	36		1.43*	1.08–1.90
Offense type		9.04*		
Rape	5		4.91**	1.64–14.68
Child molestation (extrafam.)	9		2.15*	1.11–4.16
Child molestation (incl. incest)	10		1.02	0.58–1.80
Exhibitionism	4		3.72*	1.27–10.93
Treatment participation		2.22		
Voluntary	28		1.45*	1.08–1.93
Nonvoluntary	15		1.05	0.70–1.58
Mixed	7		1.01	0.57–1.77
Treatment termination		—[a]		
Treatment completed regularly	44		1.58***	1.23–2.05
Dropped out of treatment	14		0.51***	0.39–0.67

Table 3. Continued

Variables	k	Q_{bet}	OR	CI$_{95\%}$
				Lower–upper
Methodological characteristics				
Sample size		31.43***		
Up to 50	18		4.03***	2.50–6.50
51–100	10		1.32	0.76–2.27
101–200	16		1.65**	1.13–2.41
201–500	12		1.00	0.72–1.38
More than 500	10		0.88	0.64–1.21
Maryland Scale		6.13		
Level 2 (nonequivalent)	37		1.16	0.90–1.50
Level 3 (equivalence assumed)	17		2.08***	1.40–3.08
Level 4 (matching, statistical control)	6		1.19	0.67–2.12
Level 5 (randomization)	6		1.48	0.74–2.96
Control group formation		1.64		
Treatment refused	11		1.96**	1.20–3.20
Other	47		1.37*	1.07–1.75
Source of recidivism data		4.56*		
Criminal records only	57		1.28*	1.04–1.57
Also self-report	5		3.32**	1.42–7.78

Except for the analysis on treatment approaches, studies on surgical castration are not included in the moderator analyses.

k = number of comparisons, Q_{bet} = test of between group differences (χ^2-distributed with df = number of categories − 1), OR = odds ratio; CI$_{95\%}$ = 95% confidence interval; CG = comparison group.

[a]Comparisons are based on identical CG in part; between-group differences could thus not be tested statistically.

*$P < 0.05$.

**$P < 0.01$.

***$P < 0.001$.

confounding with various content variables. For example, both hormonal medication and systemic treatment had to be subsumed under the individual category.

Offender characteristics. Programs that specifically addressed juvenile sex offenders had a higher effect than those for adult offenders. However, this difference was not significant. A related analysis showed that treatment of age-homogeneous groups tended to be more successful ($k = 48$; $r = 0.23$, $P = 0.10$).

Although the impact on specific offender groups is highly important in treatment practice, only few studies differentiated offense categories. These comparisons showed significant effects for all categories except that of intra-familial child molesting. The latter finding is related to the low recidivism baserate for incest offenders. There was a relatively large effect for rapists, but this was based on only five studies.

When sexual offenders participated voluntarily in treatment, the average *ES* was significantly positive. Obligatory participation and mixed conditions resulted in no effect. However, these differences were not significant.

Whether treatment was terminated regularly or prematurely had an impact on sexual recidivism. Whereas regular completers showed better effects than the control groups, dropouts did significantly worse. Dropping out of treatment doubled the odds of relapse and this negative effect was even homogeneous, Q (13) = 11.52, P = 0.57. In contrast, effect sizes that referred to completers revealed considerable heterogeneity, Q (43) = 100.20, P < 0.001.

Methodological characteristics. Sample size correlated significantly with effect size (r = −0.26, P = 0.03). This was particularly due to the extreme poles of the sample sizes. Comparisons based on small samples ($N \leq 50$) showed very clear effects, whereas the mean OR for very large samples was slightly below 1. This relationship could not be attributed to a publication bias only: Although in unpublished studies, the effect was somewhat less pronounced at the extreme ends of the sample size distribution, there was an even clearer linear trend compared with published studies (r = −0.34 vs. r = −0.20).

Overall, design quality did not yield a significant moderator effect. Comparisons of equivalent TG and CG (Maryland Scale Level 3 and above) revealed an average OR of 1.69 ($CI_{95\%}$: 1.26–2.28). At P = 0.06, this exceeded the OR of 1.16 for Level 2 comparisons. However, as Table 3 shows, there was no linear relationship between design quality and ES. Randomized trials also did not differ from the other comparisons, Q (1, k = 66) = 0.07, P = 0.79. Control groups containing treatment refusers revealed relatively large effects, however, these effects did not differ significantly from studies using other control groups.

The length of follow-up did not correlate with ES (r = 0.00). Different indicators of reoffending (i.e., reconviction, rearrest, etc.) also did not relate systematically to outcome variation, Q (6, k = 60) = 3.45, P = 0.49. In contrast, the sources used to gather the respective information had a significant impact on ES, Q (2, k = 62) = 7.91, P = 0.02. Comparisons using not only official records but also self-reported data had larger effects. However, this variable was confounded with the type of treatment, because all studies on hormonal medication included self-reported recidivism.

As could be expected, a higher baserate of recidivism correlated with a larger ES (r = 0.30, P = 0.01). This effect was also confounded, because informal data sources produced higher baserates. After controlling for this aspect, the relationship between baserate and ES was weaker (β = 0.23, P = 0.08).

Features of descriptive validity do not address the process of treatment and its evaluation but primarily the process of scientific reporting. Nonetheless, our analysis showed that these also related to effect size. Both the quality of documenting the treatment concept (r = 0.33, P < 0.01) and the reporting of outcome statistics (r = 0.24, P = 0.03) correlated significantly with ES.

General study characteristics. There were no significant ES differences between the various groups of countries in which the studies were performed, Q (4, k = 66) = 2.46, P = 0.65. Regarding publication type, we only found a significant effect for published comparisons (k = 40, OR = 1.62, $CI_{95\%}$: 1.23–2.13, P < 0.001). The

mean effect for unpublished comparisons was only OR = 1.14 ($k = 26$, $CI_{95\%}$: = 0.84–1.54, $P = 0.42$). However, this difference was not significant, Q (1, $k = 66$) = 2.91, $P = 0.09$.

Sensitivity analysis. The effects of moderators may be influenced strongly by a few results from studies with very large sample size (Lipsey and Wilson 2001). Therefore, we conducted a sensitivity analysis by using two different procedures. First, we excluded all comparisons with sample sizes larger than 1,000 ($k = 3$). In a second approach, all comparisons with a sample size of more than 500 were truncated to $n = 500$. With one exception, the sensitivity analyses confirmed the significant moderator effects reported in Table 3. Only the effect of unspecific offender treatment failed to reach significance when we eliminated the comparisons with sample sizes larger than 1,000 ($P = 0.08$).

Hierarchical regression. The previous analyses have repeatedly indicated problems of confounded moderators. Therefore, it is particularly relevant to see how far treatment effects are confounded with methodological and other characteristics of the evaluation. To answer this question, we computed a hierarchical regression analysis controlling sequentially for those proportions of outcome variance that could not be attributed to the treatment itself. At first, we entered unspecific and methodological study characteristics into the model. We then added offender characteristics, general treatment characteristics, and, finally, the treatment content. We entered variables that were theoretically important or empirically significant on the bivariate level ($r \geq 0.20$). At each hierarchical step, variables that did not contribute to the explanation of variance were excluded stepwise ($P > 0.10$). Because only a relatively small number of comparisons were available for the analysis, we chose this procedure in order to not overload the model with insignificant variables. Missing values were plugged with the sample mean, and analyses controlled for the effects of missing values (see Cohen and Cohen 1983). In contrast to the previous bivariate analyses, the hierarchical regression was based on a fixed effects model because we expected that the variables included would reduce a considerable part of the observed heterogeneity. The results are presented in Table 4.

As expected, heterogeneity was of moderate magnitude and not significant, Q (55, $k = 66$) = 65.40, $P = 0.16$. With 60%, the model explained a large proportion of ES variance, Q (10, $k = 66$) = 98.52, $P < 0.001$. However, one should bear in mind that due to the stepwise exclusion of variables on each cluster level, the model becomes artificially 'clean.' Only a few variables remained as independent predictors in each cluster. Methodological characteristics accounted for a considerable amount of variance (45%). Due to deficits in sample description and differentiation, offender characteristics only had a small independent impact on ES. Age homogeneity was the only relevant variable, and added 3% of explained variance. General characteristics of treatment were more important. Specificity of treatment for sex offenders, involvement of authors in the program, and a group format contributed to a 9 percentage points increase in

Table 4. Hierarchical regression.

Variable cluster	ΔR^2
Methodological characteristics	0.45***
+ Quality of outcome reporting (0.35***), Quality of treatment description (0.20*), Small sample, $N \leq 50$ (0.42***), Treatment refusers as CG (0.16†)	
− TG contains dropouts (−0.24**)	
Offenders	0.03*
+ Age homogeneity of TG (0.16*)	
General treatment characteristics	0.10***
+ Involvement of authors (0.24**), Group format (0.18†)	
− Not specific for sexual offenders (−0.19†)	
Content of treatment	0.03*
+ Cognitive orientation (0.28*)	

Changes in index direction correspond to higher (+) and lower (−) effect sizes respectively (standardized β weights are reported in brackets). Total $R^2 = 0.60$, $Q(10) = 98.52$, $P < 0.001$.
†$P < 0.10$.
*$P < 0.05$.
**$P < 0.01$.
***$P < 0.001$.

explained ES variance. Although it was only entered in the final step, a cognitive treatment orientation still added significantly to the explanation of ES variance over and above the preceding clusters. None of the other treatment variables remained in the model. Obviously, the effects of hormonal medication and behavioral conditioning methods were highly confounded with other variables. Taken together, the last two steps suggest that at least one fifth of the explained ES variance could be attributed to treatment characteristics. When interpreting this figure, we should bear in mind that this is a very conservative estimate because all other variables had been controlled already.

Discussion

Due to a recent increase in research and the multilingual approach of our review, this meta-analysis contains 80 comparisons between treatment and control groups containing a total of more than 22,000 individuals. This is currently the most comprehensive database on the outcome of sex offender treatment. Nearly one-third of the studies have been published since 2000, and approximately one-third come from countries outside of North America. These are indicators of a strong international interest in 'what works' for sex offenders. However, even though we have excluded studies containing no control group or only a comparison with dropouts, the methodological quality of the studies still remains moderate. Only 40% of the comparisons reach a level of 3 or higher on the Maryland Scale of Methodological Rigor (Sherman et al. 1997), indicating sufficient control of equivalence between TG and CG. Only seven evaluations contain a randomized

design. We have not restricted our review to these 'gold-standard' studies for the following reasons: First, limitation to a few evaluations of heterogeneous modes of treatment would not allow a differentiated analysis. Second, even randomization does not guarantee full equivalence of TG and CG (see, e.g., Marques et al. 2005). Third, as far as nonequivalence can be assessed, it tends to promote a conservative estimate of treatment efficacy (more high-risk cases in the TGs). And last but not least, the effects of our subsample of randomized studies do not differ significantly from evaluations with lower design quality (see, also, Lipsey and Wilson 1998; Lösel 1995).

Bearing the methodological problems in mind, one should draw very cautious conclusions from our meta-analysis. The most important message is an overall positive and significant effect of sex offender treatment. The mean odds ratio is 1.70 for sexual recidivism. The equivalent d coefficient of 0.29 lies within the typical range found in meta-analyses of general offender treatment (Lösel 1995; McGuire 2002). Sexual offender treatment also has an effect on general recidivism (OR = 1.67). Obviously, effective programs do not just influence sexually motivated problem behavior but also have a broader impact on criminality. This is in accordance with the experience that many sex offenders are not 'specialized' but engage in nonsexual offenses as well (Hanson and Bussière 1998). However, our analysis also shows that unspecific offender programs have no impact on sexual recidivism.

The mean rate of sexual recidivism is 11.1% in TGs and 17.5% in CGs. At first glance, this absolute difference of a little more than 6 percentage points may seem small. However, when the low baserate of sexual recidivism is taken into account, this is equivalent to a reduction of nearly 37%. For general recidivism, the reduction is 31%. Particularly in sexual recidivism, our general effect is larger than that found by Hanson et al. (2002) in their meta-analysis of psychological treatment (27%). Most probably, this is due to our inclusion of both psychological and medical modes of treatment. The average effect of physical treatment is much larger than that of psychosocial programs. The main source for this difference is a very strong effect of surgical castration, although hormonal medication also shows a relatively good outcome.

Although the very large effect of surgical castration seems to be well replicated (eight comparisons with homogeneous ES), it calls for further comment: None of the castration studies attain Level 3 on the Maryland Scale. Accordingly, we cannot assume equivalence between the TG and CG in these evaluations. Sex offenders receiving surgical castration are a highly selected and motivated group. They apply for this very intensive intervention voluntarily, whereas control individuals often refuse it or are not accepted by expert committees (e.g., Wille and Beier 1989). Hence, the TGs probably are at lower risk of reoffending than the CGs. For ethical, legal, and medical reasons, surgical castration is also rarely used in practice (Rösler and Witztum 2000). However, the very low rate of sexual recidivism in castrated offenders suggests that societies should not abandon this approach right away but perform a differentiated assessment of the pros and cons. Within an impartial and thorough process of informed consent, it may be an option

for a subgroup of high-risk sex offenders who otherwise would receive very long or lifetime detention.

Most sex offenders do not have an abnormally high level of male sex hormones (Hucker and Bain 1990; Fedoroff and Moran 1997). As with surgical castration, we must take this into account when considering the relatively strong effect of hormonal medication. Treatment with medroxy-progesteronacetate (in the US; e.g., Provera) or cyproteronacetate (in Europe; e.g., Androcur) does not seem to work by normalizing extreme testosterone levels, but by strongly reducing more or less normal levels of sexual arousal (Rösler and Witztum 2000). In addition, there are serious negative side effects that frequently lead to noncompliance and dropout (e.g., Langevin 1979). The termination of medication may rapidly increase the risk of recidivism (Meyer et al., 1992). Therefore, hormonal medication is indicated primarily for cases in which sexual arousal plays a central role in offending (e.g., Hall 1996) and should be accompanied by psychological treatment that supports compliance and has its own causal effect on sexual reoffending (Maletzky 1991; Meyer and Cole 1997).

This is why the pharmacological studies in our meta-analysis often contain psychosocial interventions as well. We have analyzed the impact of such 'treatment packages' by rating the various components separately. A regression analysis shows that only three modes of treatment have a significant impact: hormonal, behavioral, and cognitive–behavioral. The results of the first two types of program are more confounded with methodological and other study characteristics than those of the latter. After controlling for such variables, only the cognitive–behavioral orientation shows an independent treatment effect (see Table 4).

The significant positive effect of cognitive–behavioral programs is based on a solid number of 35 independent comparisons. With seven comparisons, the significant effect of classic behavior therapy has a much smaller database. The same applies to insight-oriented treatment, therapeutic communities, and other types of psychosocial programs that reveal no significant effect. That well-structured cognitive–behavioral programs work relatively well is in accordance with the literature on general offender treatment (e.g., Lösel 2001a; McGuire 2002). It is also consistent with the findings of previous reviews of sex offender treatment (e.g., Gallagher et al. 2000; Hall 1995; Hanson et al. 2002). However, the effect size for cognitive–behavioral programs in our analysis (OR = 1.45) is slightly smaller than that reported by Hanson et al. for 'current' programs that consist of mainly cognitive–behavioral approaches (OR = 1.67; direction converted by us).

Overall, we have not found that more recent programs are superior in outcome. Although treatment before the 1970s was clearly ineffective, neither programs from the 1990s nor publications after 2000 reveal stronger effects than in previous decades. Even within the cognitive–behavioral category, more current programs are not more effective than older ones. Some recent evaluations have revealed rather small or no positive effects (e.g., Friendship et al. 2003; Hanson et al. 2004; Ruddijs and Timmerman 2000; Worling and Curwen 2000). A follow-up of one of the soundest evaluations has also found no positive effect (Marques et al. 2005).

Due to the necessary follow-up lags, even recent studies may not represent all the features of the current state of the art in sexual offender treatment. One must also bear in mind that outcomes of treatment often decline when model projects are transformed into routine practice (Lösel 2001b).

The heterogeneity of outcomes within similar types of programs may be partially explained by the impact of other factors on effect size. Similar to the treatment content, these further moderators must be interpreted very cautiously: (a) Some effects are based on only a few studies. (b) The random model is less sensitive for moderator effects (Overton 1998). (c) Multiple significance testing in moderator analyses enhances the risk of an alpha error. (d) The moderators are confounded and some have no impact on the multivariate level.

Although we have found no linear relationship between design quality and outcome, there is a tendency of larger effects in studies containing equivalent treatment and control groups (at least Level 3 on the Maryland Scale). More randomized studies on the same types of treatment are needed to clarify this issue in the field of sexual offender treatment in a similar way as in other criminological areas (e.g., Weisburd et al. 2001). The impact of other methodological characteristics appears more clearly. For example, studies that include not only official recidivism data but also self-reports show larger effects. This finding is partially confounded with treatment by hormonal medication. Issues of descriptive validity such as quality of treatment description and outcome reporting are also related to larger effects.

In practical terms, the relation between sample size and treatment effectiveness is particularly important. Small studies ($N \leq 50$) reveal a large ES and large studies ($N > 500$) a small *ES*. One explanation of this result relates to publication bias. Larger samples are more likely to reveal the significance of a true small effect (Weisburd et al. 2003). Due to author or editor decisions, such large studies may be published, whereas small studies, which would have needed a larger effect size to attain significance, remain unpublished. In accordance with such an interpretation, published studies have a larger effect than unpublished studies. However, this difference is not significant, and we have found a similar – even somewhat greater – impact of sample size among the unpublished studies. Of course, this does not fully rule out some kind of publication bias, because negative results may also be less likely to be reported in unpublished studies (particularly when the researcher has a strong vested interest in the success of the program). Nevertheless, we must take a second explanation into account: In large samples, it is more difficult to maintain integrity and homogeneity of treatments or samples, and this is related to the effect size itself (Lösel and Wittmann 1989; Weisburd et al. 1993). A further finding supports the interpretation in terms of integrity: Programs in which the study authors were involved have a larger effect. As most outcome measures are beyond the influence of authors, this finding may indicate a more thorough implementation and monitoring of the program. Similar results on the effect of small samples and/ or author involvement have been observed in juvenile offender treatment (Lipsey and Wilson 1998) and developmental prevention of antisocial behavior (Farrington and Welsh 2003; Lösel and Beelmann 2003). The few comparisons for which treatment integrity can actually be rated provide further support for this relationship.

The context of treatment is also relevant for outcome. Ambulatory programs have larger effects than institutional treatment. Because we have analyzed only control group evaluations, this finding cannot be reduced to a different risk in the respective offender groups. Offender characteristics also have some impact on effect size. However, there is often a lack of detailed information on the samples. In nearly half of the studies, samples cannot even be differentiated according to the type of sexual offense. As far as this was possible, we found a significant difference in outcome. This is mainly due to a zero effect on incest child molesters resulting from the very low base rate of (official) recidivism in this group. More treatment evaluations on specific subgroups of offenders are needed to form a solid empirical basis for differential indication.

Voluntary treatment leads to a slightly better outcome than mandatory participation, and programs for adolescents are a little more effective than those for adults. Although, these differences are not significant and may be confounded with a larger baserate of recidivism in juveniles, age homogeneity is a significant moderator in the hierarchical regression. A more pronounced finding is the higher recidivism among treatment dropouts. As a consequence, studies that include dropouts in the treatment group have smaller effects. The high risk of recidivism in dropouts underlines that this group is a core problem in offender rehabilitation and controlled evaluation (Lösel 2001b). It should not only be interpreted as an individual deficit of the offender but as an interactive process and lack of fit between the program and the offender's needs and motivations (McMurran 2002). Systematic processes of program accreditation and quality management like those in Canada, England and Wales, or Scotland may help to reduce this and other problems in offender treatment. However, from a realistic perspective, we should not expect too much within a short time.

Overall, there is evidence for a positive effect of sexual offender treatment. Cognitive–behavioral and hormonal treatment are most promising. In addition, various other moderators are related to a better or worse outcome. In particular, methodological factors play an important role and seem to be confounded with treatment and offender characteristics. This problem of confounded moderators is rather general and difficult to solve (Lipsey 2003). Our hierarchical regression is only a first attempt to disentangle such patterns in the field of sexual offender treatment. We need more high-quality outcome studies that address specific subgroups of sex offenders as well as more detailed process evaluations on various treatment characteristics and components. Implementing such strategies in research and practice will further clarify 'What works for whom under which circumstances?'

Acknowledgements

Appendix: Studies integrated into the meta-analysis

Allam, J. (1998). *Effective practice in work with sex offenders: A reconviction study comparing treated and untreated offenders.* Birmingham: West Midlands Probation Service Sex Offender Unit.

Bakker, L., Hudson, S. M., Wales, D. S. & Riley, D. (1998). *"And there was Light": Evaluating the Kia Marama Treatment Programme for New Zealand sex offenders against children.* Christchurch, NZ: Psychological Service, Department of Corrections.

Barbaree, H. E. & Seto, M. C. (1998). *The ongoing follow-up of sex offenders treated at the Warkworth Sexual Behaviour Clinic.* Toronto: Centre for Addiction and Mental Health.

Barnes, J. M. (2000). *Recidivism in sex offenders: A follow-up comparison of treated and untreated sex offenders released to the community in Kentucky.* Unpublished doctoral dissertation, University of Louisville, Louisville, KY.

Berner, W. & Karlick-Bolten, E. (1986). *Verlaufsformen der Sexualkriminalität. 5-Jahres-Katamnesen bei 326 Sexualdelinquenten unter Berücksichtigung von Frühsozialisation, vorausgegangener Delinquenz, psychiatrisch-psychologischer Diagnostik und Therapie* [Developmental courses of sexual criminality]. Stuttgart: Enke.

Bluglass, R. (1980). Indecent exposure in the West Midlands. In D. West (Ed.), *Sex offenders in the criminal justice system* (pp. 171–180). Cambridge: Cambridge Institute of Criminology.

Borduin, C. M., Henggeler, S. W., Blaske, D. M. & Stein, R. J. (1990). Multisystemic treatment of adolescent sexual offenders. *International Journal of Offender Therapy and Comparative Criminology* 34, 105–113.

Borduin, C. M. & Schaeffer, C. M. (2001). Multisystemic treatment of juvenile sexual offenders: A progress report. *Journal of Psychology and Human Sexuality* 13, 25–42.

Byrne, S. M. (1999). *Treatment efficacy of a juvenile sexual offender treatment program.* Unpublished doctoral dissertation, Memorial University of Newfoundland, St. Johns, Newfoundland, Canada.

Cornu, F. (1973). *Katamnesen bei kastrierten Sittlichkeitsdelinquenten aus forensisch-psychiatrischer Sicht* [Follow-ups with castrated sex offenders]. Basel: Karger.

Craissati, J., Falla, S., McClurg, G. & Beech, A. (2002). Risk, reconviction rates and pro-offending attitudes for child molesters in a complete geographical area of London. *Journal of Sexual Aggression* 8, 22–38.

Di Fazio, R., Abracen, J. & Looman, J. (2001). Group *versus* individual treatment of sex offenders: A comparison. *Forum on Corrections Research* 13, 56–59.

Dünkel, F. & Geng, B. (1994). Rückfall und Bewährung von Karrieretätern nach Entlassung aus dem sozialtherapeutischen Behandlungsvollzug und dem Regelvollzug [Relapse and non-relapse of career offenders after release from socialtherapeutic and standard prisons]. In M. Steller, K.-P. Dahle & M. Basqué (Eds.), *Straftäterbehandlung: Argumente für eine Revitalisierung in Forschung und Praxis* (pp. 35–74). Pfaffenweiler: Centaurus.

Fedoroff, J. P., Wisner-Carlson, R., Dean, S. & Berlin, F. S. (1992). Medroxy-progesterone acetate in the treatment of paraphilic sexual disorders: Rate of relapse in paraphilic men treated in long-term group psychotherapy with or without medroxy-progesterone acetate. *Journal of Offender Rehabilitation* 18, 109–123.

Friendship, C., Mann, R. E. & Beech, A. R. (2003). Evaluation of a national prison-based treatment program for sexual offenders in England and Wales. *Journal of Interpersonal Violence* 18, 744–759.

Frisbie, L. V. (1969). *Another look at sex offenders in California* (Mental Health Research Monograph No. 12). Sacramento, CA: State of California, Department of Mental Hygiene.

Groth, N. A. (1983). Treatment of the sexual offender in a correctional institution. In J. G. Greer & I. R. Stewart (Eds.), *The sexual aggressor* (pp. 160–176). New York: Van Nostrand-Reinhold.

Guarino-Ghezzi, S. & Kimball, L. M. (1998). Juvenile sex offenders in treatment. *Corrections Management Quarterly* 2, 45–54.

Hall, G. C. N. (1995). The preliminary development of a theory-based community treatment for sexual offenders. *Professional Psychology: Research and Practice* 26, 478–483.

Hanson, R. K., Steffy, R. A. & Gauthier, R. (1992). Long-term follow-up of child molesters: Risk predictors and treatment outcome.

Hanson, R. K. & Nicholaichuk, T. P. (2000). A cautionary note regarding Nicholaichuk et al. (2000). *Sexual Abuse: A Journal of Research and Treatment* 12, 289–293.

Hanson, R. K., Broom, I. & Stephenson, M. (2004). Evaluating community sex offender treatment programs: A 12-year follow-up of 724 offenders. *Canadian Journal of Behavioural Sciences* 36, 87–96.

Hedderman, C. & Sugg, D. (1996). *Does treating sex offenders reduce offending?* (Research Findings No. 45). London: Home Office Research and Statistics Directorate.

Huot, S. J. (2002). *Recidivism, recidivism, recidivism! An update of several Minnesota recidivism studies*. Paper presented at the 21st Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers on 'Best Practices: Clinical and Research Collaborations,' Montréal, Québec, Canada. October.

Janner, J. (1959). Über die Bedeutung psychiatrischer Betreuung verurteilter Sexualdelinquenten [About the importance of psychiatric care for convicted sexual offenders]. *Schweizerische Zeitschrift für Strafrecht* 74, 310–319.

Kansas Department of Corrections. (2000). *Offender programs evaluation: Volume IV.* Topeka: Kansas Department of Corrections.

Kaul, J., Huot, S. J., Epperson, D. & Dornfeld, M. (1994). *Sex offenders released in 1988* (Unpublished report). St. Paul, MN: Minnesota Department of Corrections.

Lab, S. P., Shields, G. & Schondel, C. (1993). Research note: An evaluation of juvenile sexual offender treatment. *Crime and Delinquency* 39, 543–553.

Langevin, R., Paitich, D., Hucker, S. J., Newman, S., Ramsay, G. & Pope, S., et al. (1979). The effect of assertiveness training, Provera, and sex of therapist in the treatment of genital exhibitionism. *Journal of Behavior Therapy and Experimental Psychiatry* 10, 275–282.

Looman, J., Abracen, J. & Nicholaichuk, T. P. (2000). Recidivism among treated sexual offenders and matched controls: Data from the Regional Treatment Centre (Ontario). *Journal of Interpersonal Violence* 15, 279–290.

Lowden, K., Hetz, N., Harrison, L., Patrick, D., English, K. & Pasini-Hill, D. (2003). *Evaluation of Colorado's prison therapeutic community for sex offenders: A report of findings*. Denver, CO: Office of Research and Statistics, Division of Criminal Justice.

Maletzky, B. M. (1991). The use of medroxyprogesterone acetate to assist in the treatment of sexual offenders. *Annals of Sex Research* 4, 117–129.

Maletzky, B. M. & Field, G. (2003). The biological treatment of dangerous sexual offenders. A review and preliminary report of the Oregon pilot depo-Provera program. *Aggression and Violent Behavior* 8, 391–412.

Marques, J. K. (1999). How to answer the question, "Does sex offender treatment work?" *Journal of Interpersonal Violence* 14, 437–451.

Marshall, W. L. & Barbaree, H. E. (1988). The long-term evaluation of a behavioral treatment program for child molesters. *Behaviour Research and Therapy* 26, 499–511.

Marshall, W. L., Eccles, A. & Barbaree, H. E. (1991). The treatment of exhibitionists: A focus on sexual deviance *versus* cognitive and relationship features. *Behaviour Research and Therapy* 26, 129–135.

McConaghy, N., Blaszczynski, A. & Kidson, W. (1988). Treatment of sex offenders with imaginal desensitization and/or medroxyprogesterone. *Acta Psychiatrica Scandinavica* 77, 199–206.

McGrath, R. J., Hoke, S. E. & Vojtisek, J. E. (1998). Cognitive–behavioral treatment of sex offenders. A treatment comparison and long-term follow-up study. *Criminal Justice and Behavior* 25, 203–225.

McGrath, R. J., Cumming, G. F., Livingston, J. A. & Hoke, S. E. (2003). Outcome of a treatment program for adult sex offenders. From prison to community. *Journal of Interpersonal Violence* 18, 3–17.

McGuire, T. J. (2000). Correctional institution based sex offender treatment: A lapse behavior study. *Behavioral Sciences and the Law* 18, 57–71.

Meyer, W. J., Cole, C. & Emory, E. (1992). Depo provera treatment for sex offending behavior: An evaluation of outcome. *Bulletin of the American Academy of Psychiatry and the Law* 20, 249–259.

Mulloy, R. & Smiley, W. C. (1996). *Recidivism and treated sexual offenders*. Paper presented at the International Congress of Psychology, August 16–21, Montréal, Canada.

Nicholaichuk, T. P. (1996). Sex offender treatment priority: An illustration of the risk/need principle. *Forum on Corrections Research* 8, 38–41.

Nutbrown, V. & Stasiak, E. (1987). Research monograph: A retrospective analysis of O.C.I. cost effectiveness 1977–1981. *Ontario Correctional Institute* 2, 1–16.

Ohio Department of Rehabilitation and Correction. (1996). *Five year recidivism follow-up of sex offender releases* (Unpublished Report). Columbus, OH: Ohio Department of Rehabilitation and Correction.

Oregon Department of Corrections. (1997). *Outcome evaluation of the Jackson County Sex Offender Supervision and Treatment Program*. Retrieved 5/29/01, from http://www.doc.state.or.us/research/jackrpt2.pdf

Ortmann, R. (2002). *Sozialtherapie im Strafvollzug: Eine experimentelle Längsschnittstudie zu den Wirkungen von Strafvollzugsmaßnahmen auf Legal- und Sozialbewährung* [Social therapy in prison]. Freiburg: Edition iuscrim.

Perkins, D. (1987). A psychological treatment programme for sex offenders. In B. J. MacGurk & D. M. Thornton (Eds.), *Applying psychology to imprisonment: Theory and practice* (pp. 191–217). London: Her Majesty's Stationery Office Books.

Peters, J. J., Pedigo, J., Steg, J. & McKenna, J., Jr. (1968). Group psychotherapy of the sex offender. *Federal Probation* 32, 41–45.

Procter, E. (1996). A five-year outcome evaluation of a community-based treatment program for convicted sexual offenders run by the probation service. *Journal of Sexual Aggression* 2, 3–16.

Rasmussen, L. A. (1995). *Factors related to recidivism among juvenile sexual offenders*. Unpublished doctoral dissertation, University of Utah.

Rattenbury, F. R. (1985). *The outcomes of hospitalized and incarcerated sex offenders: A study of offender types, recidivism rates, and identifying characteristics of the repeat offender*. Unpublished doctoral dissertation, Loyola University of Chicago, Chicago, IL.

Rice, M. E., Harris, G. T. & Quinsey, V. L. (1993). Evaluating treatment programs for child

molesters. In J. Hudson & J. V. Roberts (Eds.), *Evaluating justice: Canadian policies and programs* (pp. 189–203). Toronto: Thompson.

Robinson, D. (1995). *The impact of cognitive skills training on post-release recidivism among Canadian federal offenders* (Research Report No. R-41). Ottawa, Ontario, Canada: Correctional Service of Canada.

Romero, J. J. & Williams, L. M. (1983). Group psychotherapy and intensive probation supervision with sex offenders: A comparative study. *Federal Probation* 47, 36–42.

Ruddijs, F. & Timmerman, H. (2000). The Stichting ambulante preventie projecten method: A comparative study of recidivism in first offenders in a Dutch outpatient setting. *International Journal of Offender Therapy and Comparative Criminology* 44, 725–739.

Schmid, P. (1988). *Was geschieht mit den Sexualstraftätern in der Psychiatrie? Darstellung und Bewertung der psychiatrischen Behandlung von Sexualstraftätern im Psychiatrischen Landeskrankenhaus Bad Schussenried in den Jahren 1978–1987* [What happens to sexual offenders in psychiatric clinics]. Unpublished doctoral dissertation, Universität Tübingen, Tübingen.

Smith, M. A. (2000). *The impact of prison-based drug treatment on the timing and risk of reincarceration of sexual offenders*. Unpublished dissertation, Florida State University, Tallahassee, FL.

Song, L. & Lieb, R. (1995). *Washington State sex offenders: Overview of recidivism studies* (No. 95-02-1101). Olympia, WA: Washington State Institute for Public Policy.

Stalans, L. J., Seng, M., Yarnold, P., Lavery, T. & Swartz, J. (2001). *Process and initial impact evaluation of the Cook County Adult Probation Department's Sex Offender Program: Final and summary report for the period of June, 1997 to June, 2000*. Chicago, IL: Illinois Department of Corrections.

Stürup, G. K. (1953). Sexual offenders and their treatment in Denmark and other Scandinavian countries. *International Review of Criminal Policy* 4, 1–19.

Stürup, G. K. (1968). Treatment of sexual offenders in Herstedvester Denmark: The rapists. *Acta Psychiatrica Scandinavica Supplement* 204, 1–62.

Taylor, R. (2000). *A seven-year reconviction study of HMP Grendon Therapeutic Community* (Research Findings No. 115). London: Home Office Research and Statistics Directorate.

Wille, R. & Beier, K. M. (1989). Castration in Germany. *Annals of Sex Research* 2, 103–133.

Worling, J. R. & Curwen, T. (2000). Adolescent sexual offender recidivism: Success of specialized treatment and implications for risk prediction. *Child Abuse & Neglect* 24, 965–982.

Ziethen, F. (2002). *Rückfallpräventive Effizienz der sozialtherapeutischen Behandlung von Sexualstraftätern. Evaluation der Sozialtherapie in der JVA Berlin–Tegel* [Relapse preventive efficiency of socialtherapeutic treatment for sexual offenders]. Unpublished master thesis, Freie Universität Berlin.

References

Alexander, M. A. (1999). Sexual offender treatment efficacy revisited. *Sexual Abuse: A Journal of Research and Treatment* 11, 101–116.

Aos, S., Phipps, P., Barnoski, R. & Lieb, R. (2001). *The comparative costs and benefits of programs to reduce crime. Version 4.0* (Report No. 01-05-1201). Washington: Washington State Institute for Public Policy.

Barbaree, H. (1997). Evaluating treatment efficacy with sexual offenders: The insensitivity

of recidivism studies to treatment effects. *Sexual Abuse: Journal of Research and Treatment* 9, 111–128.

Cohen, J. & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cronbach, L. J., Ambron, S. R., Dornbusch, S. M., Hess, R. D., Hornik, R. C. & Phillips, D. C., et al. (1980). *Toward a reform of program evaluation*. San Francisco, CA: Jossey Bass.

Farrington, D. P. (2003). Methodological quality standards for evaluation research. *Annals of the American Academy of Political and Social Science* 587, 49–68.

Farrington, D. P. & Petrosino, A. (2001). The Campbell Collaboration Crime and Justice Group. *Annals of the American Academy of Political and Social Science* 578, 35–49.

Farrington, D. P. & Welsh, B. C. (2003). Family-based prevention of offending: A metaanalysis. *Australian and New Zealand Journal of Criminology* 36, 127–151.

Fedoroff, J. P. & Moran, B. (1997). Myths and misconceptions about sex offenders. *Canadian Journal of Human Sexuality* 6, 263–276.

Fleiss, J. L. (1994). Measures of effect size for categorical data. In L. V. Hedges (Ed.), *The handbook of research synthesis* (pp. 245–260). New York: Russell Sage Foundation.

Friendship, C., Mann, R. E. & Beech, A. R. (2003). Evaluation of a national prison-based treatment program for sexual offenders in England and Wales. *Journal of Interpersonal Violence* 18, 744–759.

Furby, L., Weinrott, M. R. & Blackshaw, L. (1989). Sex offender recidivism: A review. *Psychological Bulletin* 105, 3–30.

Gallagher, C. A., Wilson, D. B. & MacKenzie, D. L. (2000). *A meta-analysis of the effectiveness of sex offender treatment programs*. Retrieved 5/23/2001, from http://www.wam.umd.edu/~wilsondb/papers/sexoffender.pdf.

Hall, G. C. N. (1995). Sexual offender recidivism revisited: A meta-analysis of recent treatment studies. *Journal of Consulting and Clinical Psychology* 63, 802–809.

Hall, G. C. N. (1996). *Theory-based assessment, treatment, and prevention of sexual aggression*. New York: Oxford University Press.

Hanson, R. K. & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology* 66, 348–362.

Hanson, R. K., Gordon, A., Harris, A. J. R, Marques, J. K., Murphy, W. D. & Quinsey, V. L., et al. (2002). First report of the collaborative outcome data project on the effectiveness of psychological treatment for sex offenders. *Sexual Abuse: A Journal of Research and Treatment* 14, 169–194.

Hanson, R. K., Broom, I. & Stephenson, M. (2004). Evaluating community sex offender treatment programs: A 12-year follow-up of 724 offenders. *Canadian Journal of Behavioural Sciences* 36, 87–96.

Hasselblad, V. & Hedges, L. V. (1995). Meta-analysis of screening and diagnostic tests. *Psychological Bulletin* 117, 167–178.

Hedges, L. V. & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando: Academic Press.

Hucker, S. J. & Bain, J. (1990). Androgenic hormones and sexual assault. In W. L. Marshall, D. R. Laws & H. E. Barbaree (Eds.), *Handbook of sexual assault: Issues, theories, and treatment of the offender* (pp. 93–102). New York: Plenum Press.

Langevin, R., Paitich, D., Hucker, S. J., Newman, S., Ramsay, G., Pope, S., Geller, G. & Anderson, C. (1979). The effect of assertiveness training, Provera, and sex of therapist in the treatment of genital exhibitionism. *Journal of Behavior Therapy and Experimental Psychiatry* 10, 275–282

Lipsey, M. W. (2003). Those confounded moderators in meta-analysis: Good, bad, and ugly. *Annals of the American Academy of Political and Social Science* 587, 69–81.

Lipsey, M. W. & Wilson, D. B. (1998). Effective intervention for serious juvenile offenders: A synthesis of research. In D. P. Farrington (Ed.), *Serious & violent juvenile offenders: Risk factors and successful interventions* (pp. 313–345). Thousand Oaks, CA: Sage.

Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks: Sage.

Lösel, F. (1995). The efficacy of correctional treatment: A review and synthesis of meta-evaluations. In J. McGuire (Ed.), *What works: Reducing reoffending. Guidelines from Practice and Research* (pp. 79–111). Chichester, UK: Wiley.

Lösel, F. (2000). The efficacy of sexual offender treatment: A brief review of German and international evaluations. In P. J. van Koppen & N. Roos (Eds.), *Rationality, information and progress in law and psychology. In honour of Hans Crombag* (pp. 145–170). Maastricht, NL: Metajuridica Publications.

Lösel, F. (2001a). Rehabilitation of the offender. In P. B. Baltes (Ed.), *International encyclopedia of the social & behavioral sciences* (Vol. 19, pp. 12988–12993). Amsterdam: Elsevier.

Lösel, F. (2001b). Evaluating the effectiveness of correctional programs: Bridging the gap between research and practice. In G. A. Bernfeld, D. P. Farrington & A. W. Leschied (Eds.), *Offender rehabilitation in practice. Implementing and evaluating effective programs* (pp. 67–92). Chichester, UK: Wiley.

Lösel, F. & Beelmann, A. (2003). Effects of child skills training in preventing antisocial behavior: A systematic review of randomized evaluations. *Annals of the American Academy of Political and Social Science* 587, 84–109.

Lösel, F. & Köferl, P. (1989). Evaluation research on correctional treatment in West Germany: A meta-analysis. In H. Wegener, F. Lösel & J. Haisch (Eds.), *Criminal behavior and the justice system* (pp. 334–355). New York: Springer.

Lösel, F. & Schmucker, M. (2003). *The efficacy of sex offender treatment: A brief synthesis of meta-analyses*. Report for the Campbell Crime and Justice Group (Updated). University of Erlangen–Nuremberg: Institute of Psychology.

Lösel, F. & Wittmann, W. W. (1989). The relationship of treatment integrity and intensity to outcome criteria. *New Directions for Program Evaluation* 42, 97–108.

Maletzky, B. M. (1991). *Treating the sexual offender*. Newbury Park, CA: Sage.

Marques, J. K., Wiederanders, M., Day, D. M., Nelson, C. & von Ommeren, A. (2005). Effects of a relapse prevention program on sexual recidivism: Final results from California's Sex Offender Treatment and Evaluation Project (SOTEP). *Sexual Abuse: A Journal of Research and Treatment* 17, 79–107.

Marshall, W. L. & McGuire, J. (2003). Effect sizes in the treatment of sexual offenders. *International Journal of Offender Therapy and Comparative Criminology* 47, 653–663.

McGuire, J. (2002). Integrating findings from research reviews. In J. McGuire (Ed.), *Offender rehabilitation and treatment: Effective programmes and policies to reduce re-offending* (pp. 3–38). Chichester, UK: Wiley.

McMurran, M. (Ed.). (2002). *Motivating offenders to change: A guide to enhancing engagement in therapy*. Chichester, UK: Wiley.

Meyer, W. J. & Cole, C. M. (1997). Physical and chemical castration of sex offenders: A review. *Journal of Offender Rehabilitation* 25, 1–8.

Meyer, W. J., Cole, C. & Emory, E. (1992). Depo provera treatment for sex offending behavior: An evaluation of outcome. *Bulletin of the American Academy of Psychiatry and the Law* 20, 249–259.

Overton, R. C. (1998). A comparison of fixed-effects and mixed (random-effects) models for meta-analysis tests of moderator variable effects. *Psychological Methods* 3, 354–379.

Polizzi, D. M., MacKenzie, D. L. & Hickman, L. J. (1999). What works in adult sex offender treatment? A review of prison- and non-prison-based treatment programs. *International Journal of Offender Therapy and Comparative Criminology* 43, 357–374.

Prentky, R. A., Lee, A. F. S., Knight, R. A. & Cerce, D. (1997). Recidivism rates among child molesters and rapists: A methodological analysis. *Law and Human Behavior* 21, 635–659.

Quinsey, V. L., Harris, G. T., Rice, M. E. & Lalumière, M. L. (1993). Assessing treatment efficacy in outcome studies of sex offenders. *Journal of Interpersonal Violence* 8, 512–523.

Rice, M. E. & Harris, G. T. (2003). The size and sign of treatment effects in sex offender therapy. In R. A. Prentky, E. S. Janus & M. C. Seto (Eds.). *Sexually coercive behavior: Understanding and management* (pp. 428–440). New York: New York Academy of Sciences.

Rösler, A. & Witztum, E. (2000). Pharmacotherapy of paraphilias in the next millennium. *Behavioral Sciences and the Law* 18, 43–56.

Ruddijs, F. & Timmerman, H. (2000). The Stichting ambulante preventie projecten method: A comparative study of recidivism in first offenders in a Dutch outpatient setting. *International Journal of Offender Therapy and Comparative Criminology* 44, 725–739.

Shadish, W. R., Cook, T. D. & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Sherman, L., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P. & Bushway, S. (1997). *Preventing crime: What works, what doesn't, what's promising. Report to the U.S. Congress*, Washington, DC.

Weisburd, D., Petrosino, A. & Mason, G. (1993). Design sensitivity in criminal justice experiments. In M. Tonry (Ed.), *Crime and justice* (Vol. 17, pp. 337–379). Chicago: University of Chicago Press.

Weisburd, D., Lum, C. M. & Petrosino, A. (2001). Does research design affect study outcomes in criminal justice? *Annals of the American Academy of Political and Social Science* 578, 50–70.

Weisburd, D., Lum, C. M. & Yang, S.-M. (2003). When can we conclude that treatments or programs "don't work"? *Annals of the American Academy of Political and Social Science* 587, 31–48.

White, P., Bradley, C., Ferriter, M. & Hatzipetrou, L. (1998). Managements for people with disorders of sexual preference and for convicted sexual offenders. *The Cochrane Database of Systematic Reviews*. (1998, Issue 4).

Wille, R. & Beier, K. M. (1989). Castration in Germany. *Annals of Sex Research* 2, 103–133.

Wilson, D. B. (2001). Meta-analytic methods for criminology. *Annals of the American Academy of Political and Social Science* 578, 71–89.

Worling, J. R. & Curwen, T. (2000). Adolescent sexual offender recidivism: Success of specialized treatment and implications for risk prediction. *Child Abuse & Neglect* 24, 965–982.

About the authors

Friedrich Lösel is professor of psychology and director of the Institute of Psychology at the University of Erlangen-Nuremberg, Germany. His research addresses the origins, development, prevention, and treatment of crime. He is a past president of the European Association of Psychology and Law, president of the German Criminology Society, recipient of the Sellin-Glueck Award of the American Society of Criminology, and fellow of the Academy of Experimental Criminology.

Martin Schmucker is a lecturer and research fellow at the Institute of Psychology at the University of Erlangen, Germany. His areas of research include personality disorders and crime, neuropsychological correlates of antisocial behavior, and the evaluation of treatment and prevention programs.
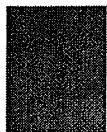
August 2007

Comparison of State Laws Authorizing Involuntary Commitment of Sexually Violent Predators: 2006 Update, Revised

BACKGROUND INFORMATION

This report updates the Institute's 2005 study, "Involuntary Commitment of Sexually Violent Predators: Comparing State Laws."[1] The 2005 report covered data through 2004; this report extends the timeframe through 2006. The report includes information on the number of residents, discharges, and program costs.

In addition to the 17 states covered in the previous report, three states have passed legislation authorizing civil commitment of sexually violent predators (SVPs). Nebraska passed a law in 2006, while New Hampshire and New York both passed laws early this year. The states with SVP laws are: Arizona, California, Florida, Illinois, Iowa, Kansas, Massachusetts, Minnesota, Missouri, Nebraska, New Hampshire, New Jersey, New York, North Dakota, Pennsylvania, South Carolina, Texas, Virginia, Washington, and Wisconsin. The Texas law is unusual as it does not require confinement; instead treatment occurs in an outpatient setting. Pennsylvania's law is only for selected sex offenders who are "aging out" of the juvenile justice system at age 21.

To collect data for this report, we asked representatives from states to complete information on a common template. Some limitations need to be acknowledged. In many cases, the state representative for the 2004 and 2006 reports differ, thus there could be different interpretations of the questions. It was particularly challenging to sort releases in relation to program staff recommendations.

---

[1] K. Gookin. (2005). *Involuntary commitment of sexually violent predators: Comparing state laws.* Olympia: Washington State Institute for Public Policy, Document No. 05-03-1101.

States do not typically track a reason for every discharge and, therefore, had to rely on best estimates. Sometimes a person is discharged for more than one reason and, therefore, numbers may not add. In terms of cost information, the full costs of a SVP law are difficult to calculate as the program's expenditures are often spread across state budgets. For example, the treatment may be supplied by a social and health services agency, legal costs paid through another agency, and security and supervision supplied by corrections. In some instances, the state representatives were unable to estimate all aspects of the costs.

Additionally, in counting individuals in various categories (revoked, recommended for release), we refined the categories from the previous report to reflect the variation in states. It is difficult, however, to capture the full nuances of each state's policies in a spreadsheet format.

FINDINGS FOR U.S.

Number of Persons Held Under SVP Laws:
4,534

Number of Persons Discharged or Released:
494; an additional 85 persons died while in custody.

Average Annual Program Costs:
$97,000 per person.

WASHINGTON STATE

Number of Persons Held Under SVP Laws:
305

Annual Program Costs:
$40.5 million in 2006

Exhibit 1

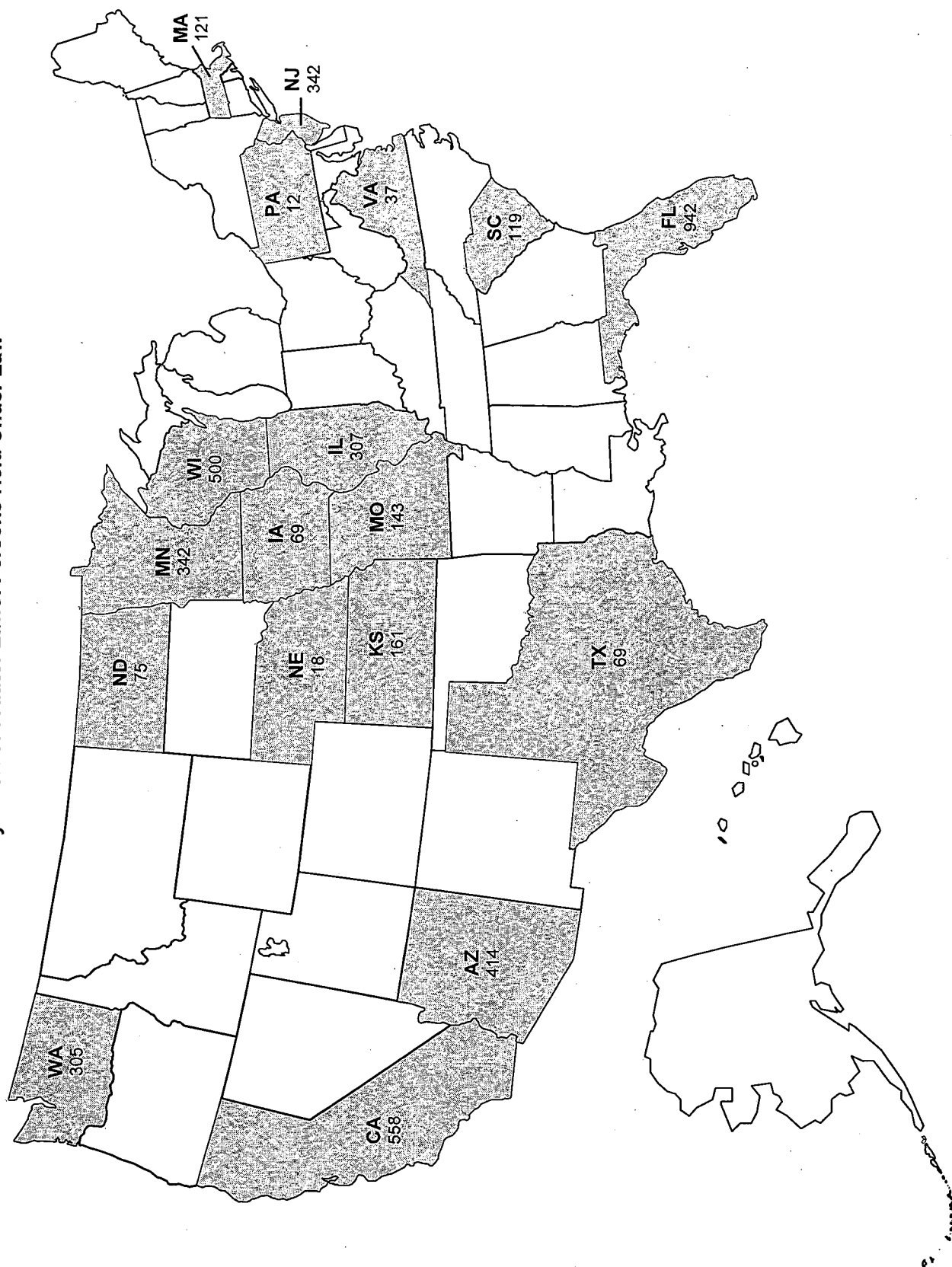**Sexually Violent Predator Laws: Persons Held Under Law**

WA 305
ND 75
MN 342
WI 500
MA 121
PA 12
NJ 342
VA 37
IL 307
IA 69
NE 18
KS 161
MO 143
SC 119
FL 942
AZ 414
CA 558
TX 69

WSIPP, 2007

Exhibit 2
2006 Program Overview

State Code (Year Law Effective)	Involved With Petition or Certification for Commitment Since Law Effective (Accumulative)	REASONS FOR DISCHARGES AND RELEASES OF PEOPLE COMMITTED					Deaths (Accumulative)	Revoked Since Release (Accumulative)*
		Discharged or Released Who Were Committed (Accumulative, Not Including Deaths)	Program Staff Recommendation	Court Decision Without Program Staff Recommendation	To Corrections (New Charges)	Other		
Arizona 36-3701 et seq. (1996)	414	87	0	86	1	0	3	13
California WIC Section 6600 et seq. (1996)	558	96	4	74	15	3	15	2
Florida ss394.910-394.931, Part V (1999)	942	28	1	16	11	0	8	9
Illinois 725 ILCS 207/1 et seq. (1988)	307	40	15	9	16	0	10	6
Iowa 229.A1 et seq. (1998)	69	15	5	10	0	0	1	1
Kansas 59-29a01 et seq. (1994)	161	16	3	0	0	12	8	0
Massachusetts Part 1 Title XVII, Ch. 123A 1 et seq. (1998)	121	5	0	5	0	0	0	N/A
Minnesota 253B.185 et seq. (1999)	342	43	27	0	3	13	2	27
Missouri 632.480 et seq. (1994)	143	10	0	0	8	2	7	0
Nebraska (2006)	18	18	14	0	4	0	0	0
New Jersey 30:4-27.24 et seq. (1994)	342	30	1	27	1	1	2	2
North Dakota 25.03-3.01 et seq. (1997)	75	20	14	1	5	0	1	0
Pennsylvania SB421 (2003)	12	N/A	N/A	N/A	N/A	N/A	N/A	N/A
South Carolina 44-48-10 et seq. (1998)	119	32	16	16	0	0	4	0
Texas Health and Safety Code 841.001 et seq. (1999)	69	N/A	N/A	N/A	N/A	N/A	1	11

| State Code (Year Law Effective) | Involved With Petition or Certification for Commitment Since Law Effective (Accumulative) | Discharged or Released Who Were Committed (Accumulative, Not Including Deaths) | REASONS FOR DISCHARGES AND RELEASES OF PEOPLE COMMITTED | | | | Deaths (Accumulative) | Revoked Since Release (Accumulative)* |
			Program Staff Recommendation	Court Decision Without Program Staff Recommendation	To Corrections (New Charges)	Other		
Virginia 37.2-900 et seq. (2003)	37	5	N/A	N/A	N/A	N/A	None	2
Washington 71.09.010 et seq. (1990)	305	19	12	1	2	2	9	4
Wisconsin 980.01 et seq. (1994)	500	30	76	8	7	2	14	23
TOTALS	4,534	494	188	253	73	35	85	100

*Unduplicated count of readmits after discharge because of a new legal process and readmits after discharge to outpatient treatment.

Notes:
Arizona: By policy, the staff does not make recommendations regarding releases.
Illinois: Recommendations for discharge/release are made by independent state evaluators, not program staff.
Massachusetts: Does not have supervised release of an SVP
Pennsylvania: Program only has "aged-out" juveniles as they turn 21 years old.
Texas: Outpatient only
Washington: One additional person is awaiting revocation.
Wisconsin: Recommends releases, but does not track the specific circumstances and their outcomes.

The following states' laws were passed in 2007 and are still being implemented.

- New Hampshire, RSA 135-E
- New York, Chapter 7 of S.3318

Exhibit 3
General Cost Information (2006)

State	SVP Resident Cost Per Year	Total Civil Commitment Budget (in millions)	Annual Salary and Benefits of Treatment Officer	DOC Inmate Cost Per Year
Arizona	$110,000	$11.3	$50,985	$20,564
California	$166,000	$147.3	N/A	$43,000
Florida	$41,845	$23.3	N/A	$19,000
Illinois	$88,000	$25.6	N/A	$21,700
Iowa	$71,000	$5.0	$56,356	$23,002
Kansas	$69,070	$10.9	N/A	$22,630
Massachusetts	$73,197	$30.7	N/A	$43,026
Minnesota	$141,255	$54.9	N/A	$29,240
Missouri	$75,920	$8.3	$25,358	$14,538
Nebraska	$93,325	$13.5	N/A	$26,031
New Jersey	$67,000	$21.9	N/A	$35,000
North Dakota	$94,728	$12.7	$55,600	$27,391
Pennsylvania	$150,000	$1.8	N/A	$32,304
South Carolina	$41,176	$2.9	N/A	$15,156
Texas	$17,391	$1.2	N/A	$15,527
Virginia	$140,000	$8.2	N/A	$23,123
Washington	$149,904	$40.5	$40,228	$29,055
Wisconsin	$102,500	$34.7	N/A	$27,600
AVERAGES	$94,017	$25.3		$25,994
	Total:	$454.7		

Notes:
New Hampshire is in the start-up phase.
New York is in the start-up phase.
Pennsylvania has only "aged-out" juveniles as they turn 21 years old. Their costs for DOC reflect 2004–2005, inflated by 3 percent.
Texas has outpatients only.

Exhibit 4
Detailed Costs

State	Treatment	Security and Supervision	Direct Care (food, clothing, supplies)	Health Care	Contracted Services	Legal Services Including Prosecutors and Defense	Psychological Evaluations	Transportation	Capital Cost/Debt Service	Administrative Overhead	Other
Arizona	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
California	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Florida	N/A	N/A	N/A	N/A	N/A	N/A	$961,263	N/A	$0	N/A	$600,000
Illinois	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Iowa	$381,169	$3,143,060	$646,650	$24,123	$24,123	N/A	N/A	$25,623	N/A	N/A	N/A
Kansas	$2,644,804	$8,039,215	$990,036	$334,628	N/A	N/A	N/A	Included in other costs	N/A	Included in other costs	$0
Massachusetts	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Minnesota	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Missouri	$6,253,000	Included in treatment costs	$605,000	$308,000	$38,510	N/A	N/A	$6,700	$0	$1,467,312	N/A
Nebraska	$1,818,066	$767,984	$222,863	Included in direct care	$0	$0	$86,007	Included in other costs	$0	$529,459	$545,456
New Jersey	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
North Dakota	$1,565,798	$125,000	$263,266	$45,375	N/A	$25,000	N/A	$5,000	$13,503	$798,953	$0
Pennsylvania	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
South Carolina	$41,176	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Texas (Outpatient only)	$143,143	$136,350	Included in contracts	N/A	$574,111	N/A	$23,384	$78,257	$0	$294,000	$0
Virginia	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Washington	$2,354,118	$10,366,000	$2,883,718	$1,249,898	$5,313,589	$5,471,222	$720,074	$23,000	$0	$809,083	$6,636,298
Wisconsin	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Notes:
Florida: Outside psychiatric evaluations are included in the daily rate. Starting in FY06–07, services were phased in beginning with clinical, medical, and other residential services, followed by internal security (December 15, 2006), and then transportation and perimeter security (April 2, 3007).
Minnesota: Will forward detailed cost information as it becomes available.
Missouri: Administrative Overhead covers staff salaries for administrative persons in the facility. Accounting, some human resources, and maintenance overhead are not included.
New York: Program not yet begun.
New Hampshire: Program not yet begun.
Pennsylvania: Plans to have detailed cost information available in 2008.
Washington: Costs are for the main facility only and do not include less restrictive facilities.

Exhibit 5
Service and Facility Providers

| State	Treatment Provider (Contractor or State)	Security Provider (Contractor or State)	Facility (Contract or State Owned)
Arizona	State	State	State
California	State (DMH)	DOC (perimeter)	State
Florida	Contract	DOC and GEO	State
Illinois	Contract	State	State
Iowa	State	State	State
Kansas	State	State	State
Massachusetts	Contract	State	State
Minnesota	State	State	State
Missouri	State	State	State
Nebraska	State	State	State
New Jersey	DHS	DOC	State
North Dakota	State	State	State
Pennsylvania	State	State	State
South Carolina	State	State	State
Texas	Contract	N/A (outpatient only)	Contract (halfway houses)
Virginia	State	State	State
Washington	State	State	State
Wisconsin	State	State	State

Note: New Hampshire and New York are still planning for implementation of their recent laws.

Exhibit 6
State Contacts

State	Name	Phone	Email
Arizona	Dr. Noggle	602-220-6482	Noggled@azdhs.gov
California	Brenda Epperly-Ellis	916-653-2088	Brenda.Epperly-Ellis@dmh.ca.gov
Florida	Teion Wells Harrison	850-921-4490	Teion_Harrison@dcf.state.fl.us
Illinois	Shan Jumper	217-322-3204 ext 5016	Shan.Jumper@illinois.gov
Iowa	Jason Smith	712-225-6948	Jsmith4@dhs.state.ia.us
Kansas	Dr. Leo Herrman		LPH0222@lsh.ks.gov
Massachusetts	Robert Murphy	508-279-8111	RFMurphy@doc.state.ma.us
Minnesota	Thomas Lundquist	218-485-5300 ext 5314	thomas.j.lundquist@state.mn.us
Missouri	Alan Blake	573-218 7079	alan.blake@dmh.mo.gov
Nebraska	Cynthia Dykeman		cynthia.dykeman@hhss.ne.gov
New Hampshire	Geoffrey Souther	603-271-8820	GSouther@dhhs.state.nh.us
New Jersey	Merrill Main		Merrill.Main@dhs.state.nj.us
New York	Rich Miraglia		cofortm@ohm.state.ny.us
North Dakota	Alex Schweitzer	701-253-3964	aschweit@state.nd.us
Pennsylvania	Dr. Vito DonGiovanni	724-675-2001	vdongiovan@state.pa.us
South Carolina	Mark Binkley		MWB86@scdmh.org
Texas	Allison Taylor	512-834-4530	Allison.Taylor@dshs.state.tx.us
Virginia	Dr. Steve Wolf	804-347-8291	steve.wolf@co.dmhmrsas.virginia.gov
Washington	Dr. Henry Richards	253-583-5933	richahj@dshs.wa.gov
Wisconsin	Dennis Doren	608-301-1455	dorendm@dhfs.state.wi.us

Document No. 07-08-1101

**Washington State
Institute for
Public Policy**
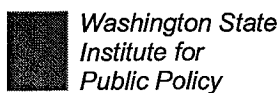
The Washington State Legislature created the Washington State Institute for Public Policy in 1983. A Board of Directors—representing the legislature, the governor, and public universities—governs the Institute and guides the development of all activities. The Institute's mission is to carry out research, at legislative direction, on issues of importance to Washington State.

HeraldNet

Everett, Wash.

Sex offender released after treatment

By _Diana Hefley_, Herald Writer

EVERETT - A 60-year-old man who the state once deemed a sexually violent predator after two rape convictions in the late 1970s was ordered set freeWednesday.

Herman "Butch" Paschke has been locked up since 1994 in the state's Special Commitment Center on McNeil Island. The state determined that Paschke was a sexually violent predator and he was civilly committed.

While locked up Paschke underwent sex offender treatment. He is the first sexually violent predator who completed the entire program, assistant Attorney General Malcolm Ross said.

Paschke hired two mental health experts who determined that he no longer met the criteria for the state to hold him as a predator. The state also hired an expert who arrived at a similar conclusion, Ross said.

That evaluation concluded that while Paschke still suffers from a mental disorder, the state lacked evidence to prove that Paschke is a danger, Ross said.

A Snohomish County judge signed an order for his release Wednesday.

Paschke is expected to reside with his wife in the 12100 block of Andrew Sater Road.

As part of the program Paschke was allowed to move home with his wife in 2005. He called authorities two days later after his wife declined to continue to be his community monitor, a condition of his release.

"She had such a bad reaction from all the media attention. She was so upset she didn't want to be his community monitor," Ross said.

Snohomish County sheriff's deputies held a community meeting about Paschke last week. He must register as a sex offender.

Reporter Diana Hefley: 425-339-3463 or hefley@heraldnet.com.
© 2009The Daily Herald Co., Everett, WA

KOMO News - Seattle, Washington

Print this article

....

One of state's most notorious sex offenders set free

by Keith Eldridge

Originally printed at http://www.komonews.com/news/local/27037919.html

PIERCE COUNTY, Wash. -- One of the state's most notorious sex offenders has just been set free. John Mathers won his freedom after spending 12 years at the state's Special Commitment Center on McNeil Island.

On Friday Mathers stood before a judge one last time just minutes before regaining his freedom. Psychiatrists concluded the 56 year old has completed his sex offender treatment and is ready for release back into the community.

"It was a long road," said John Cross, Mathers' attorney. "Mr. Mathers worked very hard. If there's going to be a success out there, he's the guy."

Mathers has been committing sex crimes since he was 20 years old when he raped two young girls and a boy. In 1980 and 1981, he twice escaped from work release and stabbed and raped two women.

After he served his prison time the state deemed Mathers too dangerous for release and put him in the Special Commitment Center in 1997.

Mathers gained notoriety in 2003 when KOMO News cameras caught Mathers, his therapist and his security escort playing golf at Fort Steilacoom public golf course in Lakewood.

The state explained this was part of easing Mathers back into society. That's the goal for all 276 sex offenders at McNeil Island.

The center was set up back in 1990 and in all that time this is only the second sex offender to be given an unconditional release.

The superintendent of the center tried to fight the unconditional release hoping to ease Mathers out into the community.

"Instead of going from 24-hour supervision to no supervision, we preferred a more ratcheting down of that supervision." said Dr. Henry Richards.

But Pierce County Superior Court Judge Stephanie Arend ordered Mathers' release.

When asked whether he could guarantee Mathers won't reoffend once he's freed, Cross said, "No one can guarantee the future."

And so out he went. On Friday was the last time Mathers will have state escorts on either side of him. Mathers now is free as long as he continues to register as a sex offender.

The Pierce County Sheriff's Office will distribute a flyer to alert the residents in Mathers' new neighborhood in Tacoma.

Court frees McNeil Island sex felon

TACOMA, Wash. -- A convicted sex offender released from the McNeil Island Special Commitment Center is now living in Tacoma.

A Pierce County Superior Court judge ordered 56-year-old John Henry Mathers released on Friday because he has completed a treatment program. He was convicted of rape and sexual assaults in the 1970s and 80s and has registered with Tacoma police as a Level 3 offender.

The Tacoma News Tribune reports that Mathers is only the second Special Commitment Center to be released in 18 years without conditions, other than registration. The other, Herman Ross Pashcke, was released a year ago in Snohomish County. The first sex offender released from McNeil Island, Joseph Aqui, violated conditions after his release in Walla Walla County. He's now in a halfway house in King County.

Information from: The News Tribune, http://www.thenewstribune.com

KIROTV.com

Man Who Admitted Raping 22 Women To Be Released From Program

Richard Thompson
KIRO 7 Eyewitness News

Posted: 6:06 pm PDT August 6, 2009Updated: 6:57 pm PDT August 6, 2009

SHELTON, Wash. -- A man who admitted to raping 22 women will become the first person unconditionally released from a program operated by the Department of Social and Health Services (DSHS).

Gary Cherry has been supervised in a program run by Washington's Department of Social and Health Services known as the Special Commitment Center on McNeil Island.

According to their Web site, the program "provides a specialized mental health treatment program for civilly committed sex offenders who have completed their prison sentences."

Cherry has been living in a Shelton house with strict rules since his prison release in 2003, but in less than a month he will have no conditions and essentially be set free.

DSHS said Cherry has completed all of his treatment and evaluations show he can no longer be labeled a "sexually violent predator" and must be released without conditions.

Local leaders said setting him free with no conditions at all will put everyone's safety at risk.

"The leash is gonna be off this wild dog that's where we're headed a leash has been taken off a wild, mad dog," said Sen. and Mason County Commissioner Tim Sheldon.

Sheriff Casey Salisbury believes Cherry has only done well living in the house in Shelton because he is under close supervision and must follow 48 conditions of release including GPS monitoring, a chaperone for all trips and periodic polygraph tests.

"If there's no supervision, he's virtually left free to do what he wants in the community and that's just not a chance I'm willing to take," Salisbury said.

KIRO 7 attempted to contact Cherry at his home, but he did not answer the door.

Cherry will be the first person in the 17-year history of the Special Commitment Center to graduate with the full support of DSHS.

His treatment has cost taxpayers more than $1.2 million.

When he is released, he'll need to register as a level-three sex offender, which means every 90 days he'll have to notify police of where he is living.

NO. 81644-1

WASHINGTON STATE SUPREME COURT

In Re the Detention of:

David W. McCuistion

        Petitioner-Appellant.

DECLARATION OF
SERVICE

I, Jennifer Dugar, declare as follows:

On this 13th day of October, I sent via FedEx Overnight true and correct cop(ies) of Statement of Additional Authorities and Declaration of Service, postage affixed, addressed as follows:

David Donnan
Washington Appellate Project
1511 3rd Ave, Ste 701
Seattle, WA 98101

Nancy Collins
Washington Appellate Project
1511 3rd Avenue, Ste. 701
Seattle, WA 98101

I declare under penalty of perjury under the laws of the State of Washington that the foregoing is true and correct.

DATED this 13th day of October, 2009, at Seattle, Washington.

JENNIFER DUGAR